IFAC

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

WARSZAWA 1969

Bionics

Fourth Congress of the International Federation of Automatic Control
Warszawa 16–21 June 1969

TECHNICAL SESSION 70

NOT

Organized by
Naczelna Organizacja Techniczna w Polsce

INTERNATIONAL FEDERATION OF AUTOMATIC CONTROL

Bionics

TECHNICAL SESSION No 70

FOURTH CONGRESS OF THE INTERNATIONAL
FEDERATION OF AUTOMATIC CONTROL
WARSZAWA 16 – 21 JUNE 1969

Organized by
Naczelna Organizacja Techniczna w Polsce

Contents

Wydawnictwa Czasopism Technicznych NOT - Polska

Zakład Poligraficzny WCT NOT. Zam. 78/69.

THE FUNCTIONAL CONTROL OF THE EYE-TRACKING-SYSTEM AND ITS DIGITAL SIMULATION

Gerhard Vossius and Jürgen Werner
J.W.Goethe-Universität, Frankfurt/Main, Germany

1. The functional organization of the eye-tracking-system

The eye-tracking-system can be divided according to fig. 1 into the retina as the measuring device and into a continuously and a discontinuously (saccadicly) operating system, attributed to the higher centers of the brain. These higher centers are connected with the eye muscle nuclei together with the fasciculus longitudinalis medialis and the eye muscles which are attached to the eye-balls. Within the eye musculature there are the muscle spindles and the tendon organs. It is assumed that the feedback pathway of these spindles and tendon organs leads to the eye muscle nuclei. By this, the existence of a second, inner control-loop is suggested, the transfer-function of which has been discussed in former papers.

The investigations of the last years have been concentrated to the experimental analysis and the mathematical formulation of the mechanisms of the saccadic and the continuous system of the higher centers. The different band-width of the discontinuously and the continuously operating branches permits to separate the electrically measured eye-movement into these two components. By this, it is possible to add one of them or both in the positive or negative sense with a variable gain to the target-signal, so that the system may be tested under various conditions. A part of the experimental investigation has been carried out together with L. Goodman and G. Bowman, Bioengineering Group, Case Institute of Technology, Cleveland.

In the following we want to illustrate the properties of the saccadic system and then we shall comment upon a general mathematical model, which is appropriate to the manifold properties of identification and prediction of voluntary movement control.

2.) The saccadic system and its optimization

The saccadic system has been described earlier as a sampled data system [11]. But this model has to be modified according to our experiments, as far as the gain of the control system is not constant, as assumed by Young and Stark. Introducing into such a control system (fig. 3) an external feedback K_r, we have the following error f_n and amplitude a_n in the n-th sampling period:

$$f_n = s - a_n (1 - K_r)$$

$$a_n = a_{n-1} + f_{n-1} K_{vn}$$

or after substitution:

$$f_n = f_{n-1} \cdot \{ 1 - K_{vn} (1 - K_r) \}$$

In order to obtain the optimal gain K_{vn} the error f_n must be set to zero:

$$K_{vn\ opt} = \frac{1}{1 - K_r}$$

We see, that the optimal internal gain for the n-th sampling period is independant of the sequence of the preceding values. It depends only on the external feedback K_r. In principle, such an optimization may be done in one step. Regarding our experiments, the saccadic system of the eye-movement is in fact capable of identifying the sign and the amount of the external feedback up to a certain extent and of adapting the inner gain K_{vn} to the new conditions, compensating the external feedback. Even with a negative feedback of $K_r = -10$ the optimal value is reached after a few steps. In fig. 5 we recognize the remarkable fact, that the system often identifies the new feedback conditions already after the first step so well, that the desired value $K_{vn\ opt.}$ is nearly reached. After this K_{vn} changes in an oscillating manner before attaining the final value. It is not surprising, that this final value is not reached directly, because even under normal conditions without any external feedback, there is generally no unity gain as should be expected. So we can distinguish an intended and a real gain, evidently caused by an internal noise level:

$$K_{vn\ eff} = K_{vn\ int} \pm \varepsilon_n$$

That means, that the organism needs a certain time to separate the influence of the feedback and of the own noise-level.

Such a hypothesis is supported by the fact, that the final value $K_{vn\ opt.}$ is reached very often immediately after a relatively large deviation K_{vn-1}, that is a large signal-to-noise-ratio.

The capability to identify changes of the feedback directly enables the organism by way of controlling his own functional state, to compensate interferences e.g. of other parts of the central nervous system.

In another series of experiments under normal conditions we have shown that the saccadic system is capable to identify in an extremely short time interval the structure of deterministic input-signals e.g. intricate sequences of step-functions with a certain periodic structure, and by this, to predict the signal in the course of the experiment. In this case, too, identification is achieved in a nearly optimal time.

3. A general model of the predictor-system

The experiments of the last years have shown, that one of the most essential properties of the control-systems of voluntary movement is the capability of prediction. Manifold experiments with continuous and sampled input-signals yielded a predictor-system, which is able to compensate dead-time, to bridge signal-gaps or continue interrupted signals. The prediction-interval depends on the signal-structure as well as on the length of the period, during which information is accumulated and used for the signal-synthesis. It has been shown in former papers [6,7,8] that these biological control-mechanisms do neither use frequency-analysis nor statistical analysis. Nor is the aim of optimization a minimum of the quadratic error. Wiener-filters e.g. are not adequate models for such predictor-systems. The system to be developed should analyse the signals and their deterministic structure in the time-domain and should generate a prediction for a time interval adapted to the amount of the relevant information. Based on this concept, we aimed at a generalization of the observed phenomena which would permit not only a better insight into the functional principles realized in the eye- and hand-tracking-system, but which could also serve as a basis for the development of models for other

data-processing technical, but particularly biological systems with predictive properties. As the manifold possibilities of the organism cannot be described by only one actually known theoretical principle, we choose a system with parallel branches (fig. 6), which can work autonomously up to a certain extent, but contribute with different weights to the synthesis of the output-signal. The weight of the parallel components is adapted to the structure of the input-signals. Using this model, we obtain a complete system which certainly does not satisfy the often extremely economic aspects of technical systems, but which is typical for biological control-systems, often disposing of possibilities of variations in order to be capable of adaptation and compensation.

We should make a difference between predictor-systems based on extrapolation and on reproduction. We understand by extrapolation a prognosis of a series of events f(t) with the knowledge of the function f(t) and its derivatives $\frac{d^{\nu} f(t)}{dt^{\nu}}$ during a certain interval, whereas prediction based on reproduction presumes an identification of repetitive features of the function f(t). Not all biological systems, which are able to predict by way of reproduction, are capable to reproduce input-signals for any time even without any further stimulation. Therefore we divided such predictive properties into a "long-time-prediction" and a "pattern-reproduction". The first is mainly based on an unconscious signal-analysis, whereas the latter is essentially a conscious data-storing and -processing procedure. For certain biological models it will also be useful to make a difference between "short-time-extrapolation" and "long-time-extrapolation".

The following statements will be formulated very generally, so that the functional principles may be recognized and applied, with no restrictions by the special biological problems.

For the mathematical description of the input-functions we choose the following form:

$$f(t) = \sum_{n=1}^{N^*} \int_{t_{n-1}}^{t_n} \varphi_n(\tau) \cdot \delta(\tau - t)\, d\tau$$

i.e. $f(t) = \varphi_n(t)$ for $t_{n-1} \leq t < t_n$, $f(t) = 0$ else

according to the definition of the Dirac - δ-function. The function f is known up to the time $t = t^*$. Discontinuities of the function and its derivatives occur only in finite time-intervals $\Delta t_n = t_n - t_{n-1}$.

3.1. Extrapolation

3.1.1. "Short-time-extrapolation"

A rather elementary extrapolation-system uses only the information available at $t = t^*$. The extrapolation is performed for $t = t^* + \Delta\tau$. In order to get a satisfying extrapolation, we must require $t^* + \Delta\tau < t_n^*$, presuming that the functions and its derivatives are continuous during the interval $t_{n-1}^* < t < t_n^*$. The system cannot predict discontinuities. In the interval $t_{n-1}^* < t < t_n^*$ the function is expanded into a Taylor-series:

$$f(t) = \sum_{\nu=0}^{N} \frac{(t - t^*)^\nu}{\nu!} \cdot f^{(\nu)}(t^*) \qquad N \to \infty$$

For $t = t^* + \Delta t$ we obtain the following extrapolation:

$$f(t^* + \Delta\tau) = \sum_{\nu=0}^{N} \frac{f^{(\nu)}(t^*)}{\nu!} \Delta\tau^\nu \qquad N \to \infty$$

Another extrapolation-system, comprising the described elementary one as a special case, is more appropriate to the biological facts. It processes all the information disposible during the interval $t^* - \Delta t < t < t^*$. The condition $t^* - \Delta t > t_{n-1}^*$ should be fulfilled for an efficacious extrapolation. All extrapolations carried out during the interval Δt are weighted by a function $g(\tau)$ and form as a whole the output-signal for $t = t^* + \Delta\tau$:

$$f(t^*, \Delta t, \Delta\tau) = \int_{\Delta t}^{0} \sum_{\nu=0}^{N} \frac{f^{(\nu)}(t^* - \tau)}{\nu!} (\Delta\tau + \tau)^\nu \cdot g(\tau)\, d\tau$$

The weighting function may e.g. have, according to physiological facts, the following form:

$$g(\tau) = \frac{e^{(\Delta t-\tau)/T} - 1}{-T\left(e^{\Delta t/T} - 1\right) + \Delta t}$$

The parameter T permits an individual adaptation of the function, whereas the expression in the denominator guarantees, that the following condition will be satisfied:

$$\int_{\Delta t}^{0} g(\tau)\, d\tau = 1$$

When weighting at descrete moments $t = t^* - \Delta t$, $t = t^* - 2\cdot\Delta t$, etc., we obtain the following extrapolation:

$$f(t^*+\Delta\tau) = \sum_{\mu=0}^{M} \sum_{\nu=0}^{N} \frac{f^{(\nu)}(t^*-\mu\cdot\Delta t)}{\nu!} (\Delta\tau + \mu\cdot\Delta t)^{\nu} \cdot g_{\mu}$$

with a similar, but discrete weighting-function g_{μ}. This weighted extrapolation generates even for signals, which have not an extremely regular structure a good extrapolation for short intervals Δt and $\Delta\tau$.

3.1.2. "Long-Time-Extrapolation"

The systems discussed up to now use only the information available in a relatively short time interval. The following procedures take into account a larger part of the function ($t^*_{n-1} \leqslant t < t^*$) and permit generally a longer prediction-interval ($t^* < t < t_n$).

The following function is considered and used for the extrapolation:

$$f(t) = \int_{t^*_{n-1}}^{t^*} \varphi^*(\tau)\, \delta(\tau - t)\, d\tau \qquad t^*_{n-1} \leq \tau < t^*$$

The function $\varphi^*(\tau)$ can be approximated by a polynomial of m^*-th order:

$$\varphi^*(\tau) = \sum_{m=0}^{m^*} a_{n^*,m} \cdot \tau^m$$

In the technical realization the coefficients $a_{n^*,m}$ may be calculated using the method of least squares fit. The extrapolation itself consists of the simple extension of the validity of φ^*:

With $\varphi_{n^*}(\tau) = \varphi^*(\tau) \qquad t^*_{n-1} \leq \tau < t^*$

the extrapolated function will be:

$$f(t) = \int_{t_{n-1}}^{t_n^*} \varphi_n^*(\tau) \cdot \delta(\tau - t)\, d\tau$$

The system is appropriate for deterministic signals disturbed by noise, but will not predict discontinuities, neither.

3.2. Prediction by reproduction

3.2.1. "Long-Time-Prediction"

The so-called "long-time-prediction-system" is able to identify all sorts of periodicity and by this to predict a function e.g. up to the next characteristic point (extremum, discontinuity etc.). The identification of periodicity may be achieved by the following procedure. In the following equations the index λ is increased up to λ^*, until the sum of differences is zero. If λ^* exceeds a certain amount, the procedure is stopped without success:

With $\lambda = 1,2,3 \cdots \lambda^*$ and $\Delta t_n = t_n - t_{n-1}$

$$\sum_{\mu=1}^{\lambda} \Delta t_{n^*-\mu+1} - \Delta t_{n^*-\mu-\lambda+1} \doteq 0$$

$$\sum_{\mu=1}^{\lambda} |\varphi_{n^*-\mu+1}| - |\varphi_{n^*-\mu-\lambda+1}| \doteq 0$$

With the conditions fulfilled by λ^*, the function may be predicted up to the next characteristic point:

$$\Delta t_{n^*+1} = \Delta t_{n^*-2\lambda^*+1} + \varepsilon_\tau$$

$$\varphi_{n^*+1} = -\varphi_{n^*-2\lambda^*+1} + \varepsilon_\varphi$$

with

$$\varepsilon_\tau = \varepsilon_\tau(\lambda^*, L, T_p)$$
$$\varepsilon_\varphi = \varepsilon_\varphi(\lambda^*, L, T_p)$$
$$T_p = t_{n^*} - t_{n^*-2\lambda^*}$$

During a sequence of executions of this procedure it will occur, that the same λ^* is reached L times without interruption. The functions ε_τ and ε_φ have to be individually determined for each system. They take into account, that an efficacious prediction may occur after a time of learning. The course of the learning-phase is influenced by the parameters λ^*, L and T_p. Gradually ε_τ and ε_φ approach zero, ε_τ sometimes will even be negative.

3.2.2. "Pattern-Reproduction"

After the end of the learning phase ε_τ and ε_φ change only negligibly, and an effective pattern-reproduction becomes possible. For κ^* periods we obtain the following prediction by way of reproduction: ($\varepsilon_\tau = 0$, $\varepsilon_\varphi = 0$).

$$f(t) = \sum_{k=1}^{\kappa^*} \sum_{\lambda=1}^{2\lambda^*} \int_{t_r}^{t_{r+1}} \varphi_{r+1}(\tau) \cdot \delta(\tau - t)\, d\tau \quad ; \quad r = n^* + \lambda - 1 + 2\lambda^*(k-1)$$

and analogous to the procedure of the preceding chapter:

$$\left.\begin{aligned} \Delta t_{n^*+i} &= \Delta t_{n^*-2\lambda^*+i} \\ \varphi_{n^*+i} &= \varphi_{n^*-2\lambda^*+i} \end{aligned}\right\} \quad 1 \leq i \leq 2\lambda^*$$

3.3. The coordinating system

The described parts of the system contribute in a different way to the synthesis of the output-signal. The weight of each branch is determined by a system called "signal-identification" (fig. 6). The procedure is not a statistical analysis, but a classification with simple criterions, e.g. the frequency of extrema and changes of direction of the signals and its derivatives, which permits a rough estimation of the signal-structure. The decisions are controlled according to the achieved performance, a measure of which is the momentaneous error. By this, the weight of each branch may be altered. If the error or the signal-velocity exceeds certain limits, corrective steps are generated by the saccadic system.

The system, as a whole, is able to adapt the data-processing parts to different classes of input-signals. By this, it can dispense with all statistical error-criterions. Surely, it is especially appropriate to deterministic signals. The inevitable data-processing-time involves limitations of performance.

4. Simulation of the eye-tracking-system

The applicability of this mathematical concept has been checked by the simulation of the eye-tracking system. The program for a digital computer comprises the possibilities of prediction by way of extrapolation and reproduction, coordinated with those of the error-correction. In contrast

to the hand-tracking-system there is no "pattern-reproduction". We determined the functions ε_τ and ε_φ experimentally and investigated the factors influencing the saccadic systems. These are particularly the momentaneous error F_x, the error of velocity $F_{\dot{x}}$, the signal-velocity $\dot{x}$ itself and a factor σ characterizing the irregularity of the signal.

In the realized classification scheme all signals with a first resp. second derivative disappearing during fixed intervals are classified as type-1-and 2-signals. The remaining part are type-3-signals. According to this classification certain predictor-mechanisms are preferred, e.g. type-3-signals are processed by the short- and long-time-extrapolation-system and type-2-signals by the short-time-extrapolation and the long-time-predictor-system. Processing type-1-signals long-time-prediction is prevailing. Further modification of weighting depends on the estimation of performance. Some of the results of the simulation are shown in comparison to those of the experiment in fig, 8 - 11.

The whole simulation of this special system has been derived from the previously developed functional concept. By this, a distinct understanding of the organization of eye-movement-control was achieved. Beyond this, one can expect that such a concept of identification and prediction should be found in other biological systems, as well.

Literature

1) Gibson, J. E.: Nonlinear Automatic Control New York, 1963

2,3) Robinson, D. A.: The mechanics of human saccadic and smooth eye-movement, J.Physiol. 174 (1964), 245 and 180 (1965), 568.

4) Stark, L., Vossius, G. and Young, L.R.: Predictive control of eye-tracking movement, Trans. IRE, Human Factors in Electr. HFE-3,52.

5) Schlitt, H.: Systemtheorie für regellose Vorgänge, Berlin, 1960.

6) Vossius, G.: Die Regelbewegung des Auges, in: Aufnahme und Verarbeitung von Nachrichten durch Organismen, Stuttgart, 1961.

7) Vossius, G.: Die Vorhersageeigenschaften des Systems der Willkürbewegung, in Neuere Ergebnisse der Kybernetik, München, 1964.

8) Vossius, G.: Der kybernetische Aspekt der Willkürbewegung, in Progress in Biocybernetics, Amsterdam, London, New York, 1965.

9) Vossius, G., Goodman, L. and Bowman, G.: Die Funktion der kontinuierlichen und sakkadischen Komponente der Augenbewegung, Pflügers Archiv ges. Physiol.,(1966), 289, R 65.

10) Wiener, N.: Extrapolation, Interpolation and Smoothing of stationary Time-Series, New York, 1950

11) Young, L. R., Stark, L.: Ein Abtastmodell der Augenfolgebewegung, Regelungstechnik 11 (1963), 148

Figures:

Fig.1: Schematic model of functional organization of the eye-tracking-system (modified after VOSSIUS 1960/61)

Fig.2: Experimental setup

Fig.3: The saccadic system as a sampled data system

Fig.4: Saccadic eye movement with experimental negative feedback $K_R = -3.0$

Fig.5: Adaptation of the gain K_v

Fig.6: The general functional concept

Fig.7: Target-function

Fig.8 and 9: Eye-movement in the experiment

Fig.10 and 11: Eye-movement in the simulation

Explications to figures

fig.1	Auge	=	eye
	Netzhaut	=	retina
	Kont. System	=	continuous system
	Diskont. System	=	discontinuous system
	Vestibularis	=	vestibularis
	AMK. M. HL.	=	eye muscle nuclei with fasciculus longitudinalis medialis
	AM.	=	eye muscles
	PRZ.	=	muscle spindles and tendon organs
	H. Z.	=	higher centers of CNS
fig.2	Schirm	=	screen
	Vorgabe (signal)	=	target (signal)
	Lampe	=	lamp
	Spiegelgalvanometer	=	galvanometer
	Auge	=	eye
	Vorverstärker	=	amplifier
	Galvanometer-Verstärker	=	galvanometer-amplifier
	Solarzelle	=	photo-cell
	Analogrechner	=	analogcomputer
	Augenfolgebewegung	=	eyemovement
	Sakkaden	=	saccadic movement
	Kont.	=	continuous movement
	Fehler	=	error
	Registrierung	=	recorder
	Frequenzgenerator	=	frequency-generator
fig.3	----------		
fig.4	Vorgabe	=	target
	Auge	=	eye
	Sakkaden	=	saccadic movement
	Kont.	=	continuous movement
	Schaltsignal	=	differentiated signal
fig.5	----------		

fig. 6	Signalidentifikation	=	signal identification
	Gewichtung	=	weighting
	Kurzzeitextrapolation	=	short-time-extrapolation
	Langzeitextrapolation	=	long-time-extrapolation
	Musterreproduktion	=	pattern-reproduction
	Fehlerkorrektur	=	error-correction
	Erfolgsanalyse	=	estimation of performance
fig. 7	----------		
fig.8-11	Vorgabe	=	target
	Folge	=	eye tracking movement

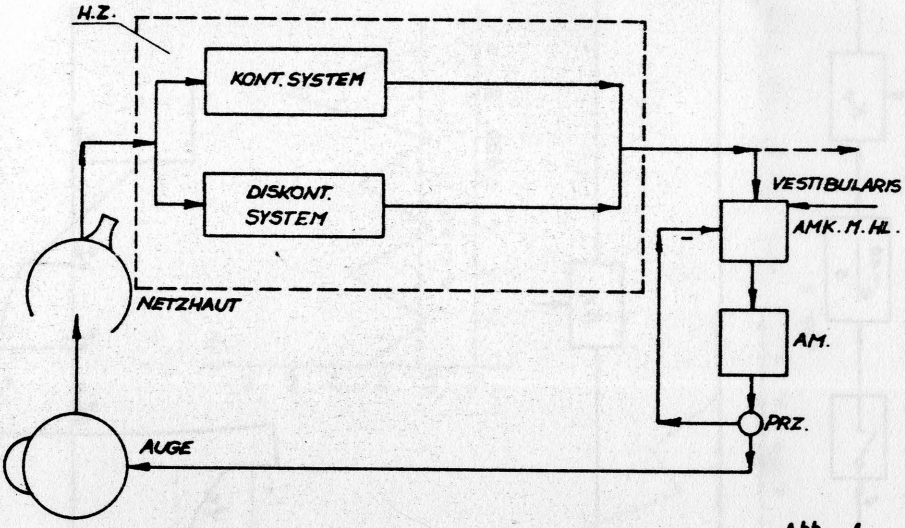

H.Z.
KONT. SYSTEM
DISKONT. SYSTEM
VESTIBULARIS
AMK. M. HL.
AM.
PRZ.
NETZHAUT
AUGE

Abb. 1

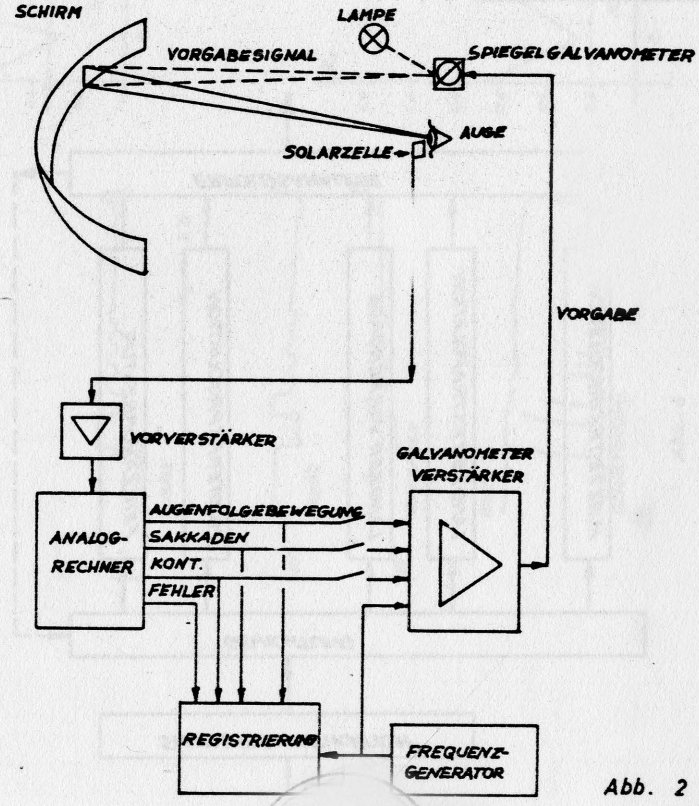

SCHIRM
LAMPE
VORGABESIGNAL
SPIEGELGALVANOMETER
AUGE
SOLARZELLE
VORGABE
VORVERSTÄRKER
GALVANOMETER VERSTÄRKER
ANALOG-RECHNER
AUGENFOLGEBEWEGUNG
SAKKADEN
KONT.
FEHLER
REGISTRIERUNG
FREQUENZ-GENERATOR

Abb. 2

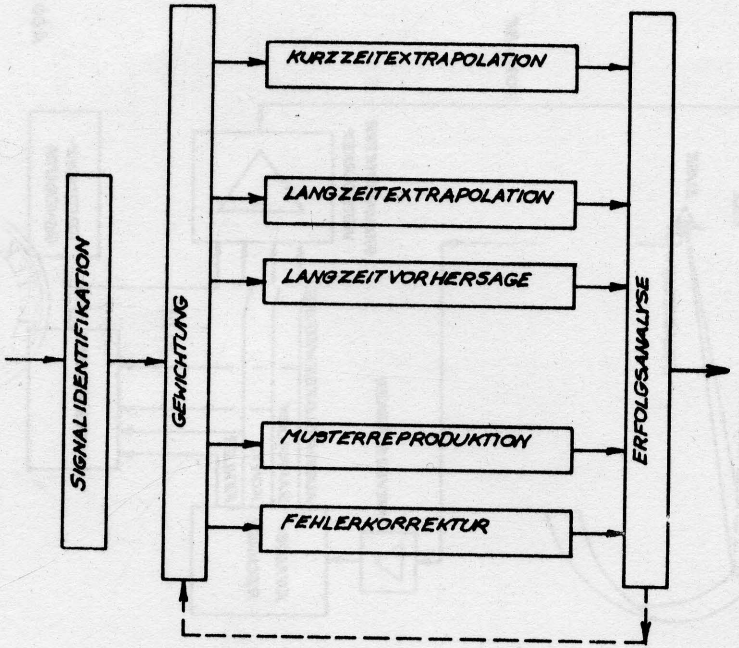


Abb. 6

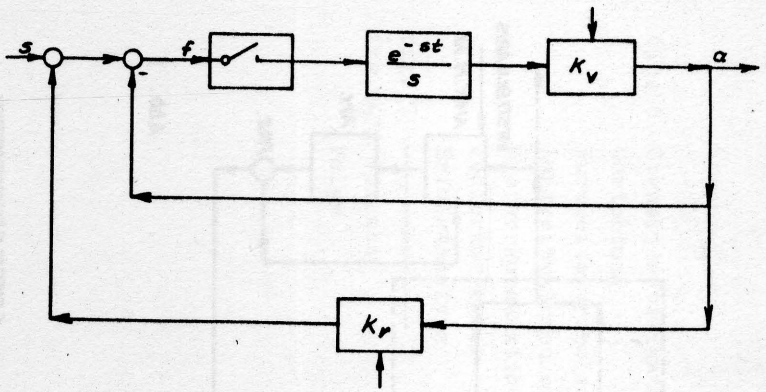


Abb. 3

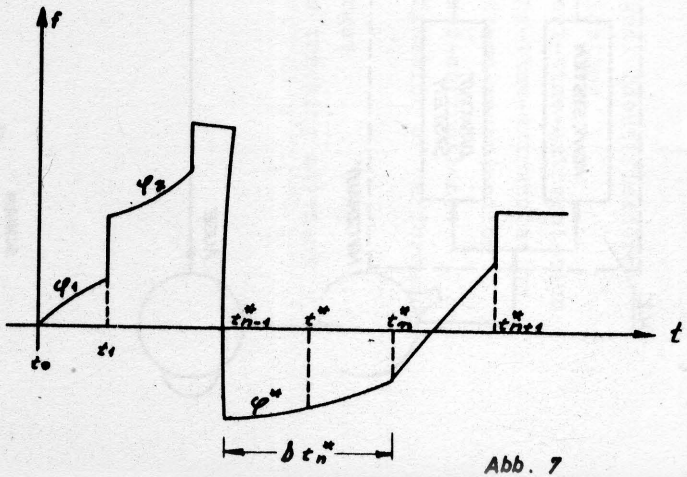


Abb. 7

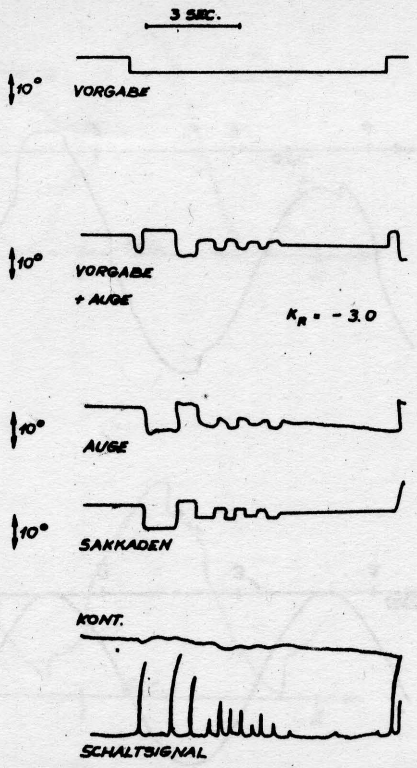

3 SEC.
10°
VORGABE
10°
VORGABE
+ AUGE
K_R = - 3.0
10°
AUGE
10°
SAKKADEN
KONT.
SCHALTSIGNAL

Abb. 4

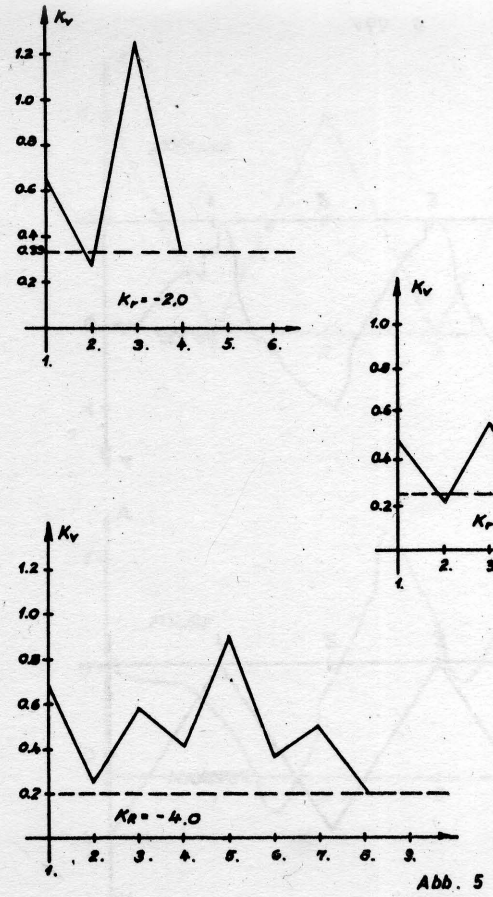
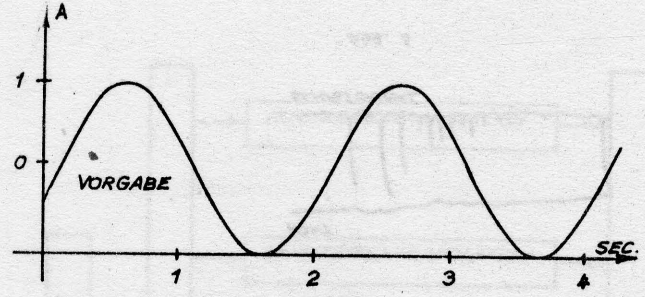

K_V
1.2
1.0
0.8
0.6
0.4
0.33
0.2
K_R = -2,0
1. 2. 3. 4. 5. 6.
K_V
1.0
0.8
0.6
0.4
0.2
K_R = - 3.0
1. 2. 3. 4. 5. 6. 7. 8.
K_V
1.2
1.0
0.8
0.6
0.4
0.2
K_R = - 4.0
1. 2. 3. 4. 5. 6. 7. 8. 9.

Abb. 5


A
1
0
VORGABE
SEC.
1
2
3
4

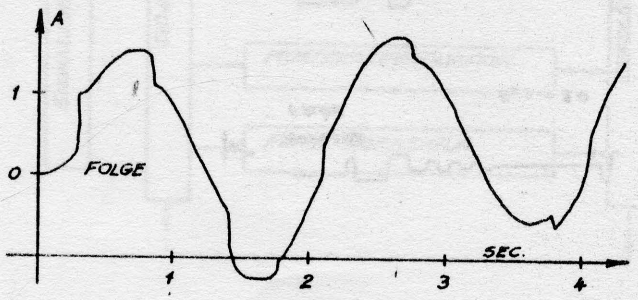
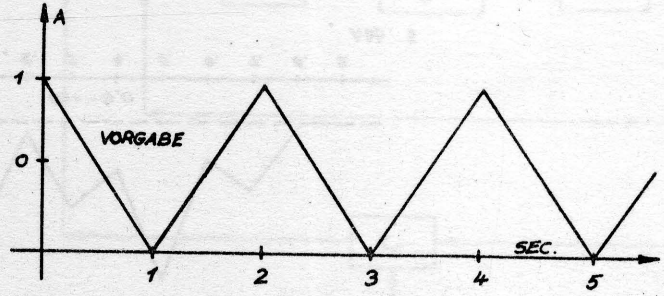

A
1
0
FOLGE
SEC.
1
2
3
4

Abb. 8


A
1
0
VORGABE
SEC.
1
2
3
4
5

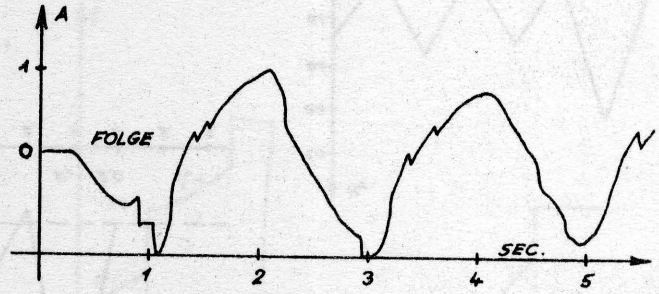
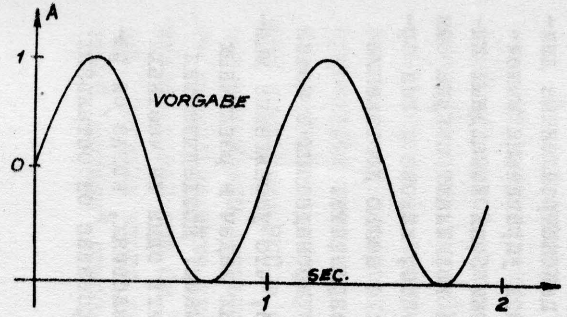

A
1
0
FOLGE
SEC.
1
2
3
4
5

Abb. 9


A
1
0
VORGABE
SEC.
1
2

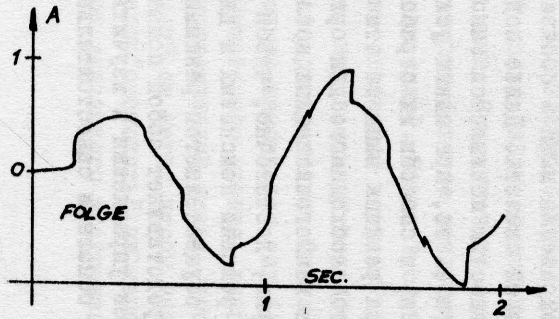
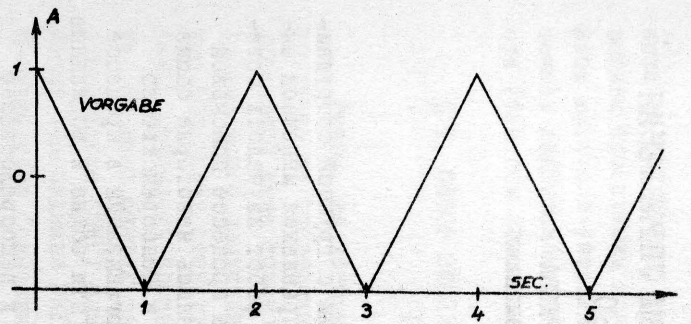

A
1
0
FOLGE
SEC.
1
2

Abb. 10


A
1
0
VORGABE
SEC.
1
2
3
4
5

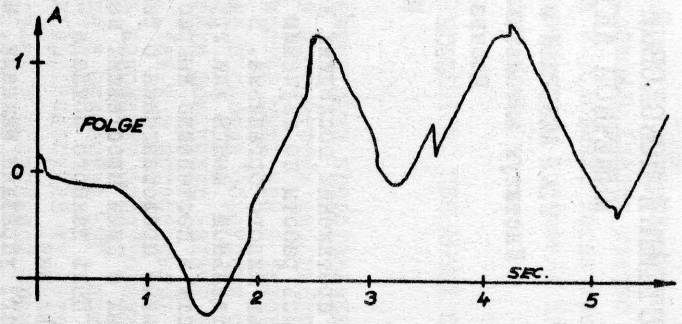

A
1
0
FOLGE
SEC.
1
2
3
4
5

Abb. 11

ПРОСТЕЙШИЙ ПОИСКОВЫЙ МЕХАНИЗМ УПРАВЛЕНИЯ МЫШЕЧНОЙ АКТИВНОСТЬЮ

М.А. АЙЗЕРМАН и Е.А. АНДРЕЕВА
Институт автоматики и телемеханики
Москва
СССР

На протяжении последних лет авторами с группой сотрудников велись работы по изучению проблем управления мышечной активностью живого организма. В ходе этих работ изучалось поведение отдельной мышцы или группы мышц в процессе поисковой активности и постепенно выкристализовывались некоторые общие и модельные представления о работе мышц и непосредственно связанных с ними простейших нейронных организаций в процессе поиска. Цель данного доклада - изложить эти общие и модельные представления.

Мозг, управляя мышцами, направляет к непосредственно связанным с ними нервным организациям (спинальным сегментам) командные залпы импульсов и под действием этих командных импульсов и импульсных сигналов, поступающих на спинальный сегмент от рецепторов самой мышцы, осуществляются целенаправленные сокращения мышц и соответственно целенаправленные движения костей скелета. Далее под механизмом управления мышечной активностью понимается закон выработки этих командных импульсов (т.е. то, при каких условиях они возникают и куда они подводятся) и способы их отработки (т.е. то, какова и как организуется реакция мышц на эти импульсы) с целью достижения некоторого необходимого для организма результата.

Даже немногочисленных исследований, проведенных в нашей лаборатории, достаточно, чтобы убедиться, что управление мышцами при решении поисковых и двигательных задач в различных условиях осуществляется разными механизмами. Докладываемая работа представляет собой попытку выделить один из частных механизмов управления и изучить его в условиях, когда он работает в основном самостоятельно, изолированно от остальных

механизмов. Именно изучается механизм, который мы условно называем простейшим поисковым механизмом (далее кратко ППМ). Этот механизм управляет мышцами в условиях, когда требуется точное поддержание некоторого найденного положения или скорости (частоты колебаний) суставных углов.

§ I. ОБЗОР ЭКСПЕРИМЕНТАЛЬНЫХ ФАКТОВ, СВЯЗАННЫХ С РАБОТОЙ ППМ

Факты, на которых основываются наши представления о ППМ, получены из проведенных в нашей лаборатории наблюдений над указанной выше поисковой активностью мышц как в условиях искусственного замыкания организма внешними обратными связями, что позволяет ставить мозгу четкую поисковую задачу и постепенно наблюдать процесс поиска, так и в естественных условиях. Во всех опытах записывалась накожными электродами электромиограмма (ЭМГ) мышцы или мышц и огибающая ЭМГ, полученная после прохождения ЭМГ через фильтр, состоящий из детектора, инерционного (одноемкостного) звена с постоянной времени T и катодного повторителя.

I°. Опыты по поддержанию минимума болевого раздражения, когда оно зависит от напряженности мышц.

В этих опытах I стимул, вызывающий болевое раздражение за счет искусственной обратной связи, зависел только от огибающей ЭМГ одной или нескольких мышц. Существовало единственное значение огибающих ЭМГ, при которых болевое раздражение было минимальным (в частности, равным нулю). Тогда в случае одной мышцы I в плоскости "значение огибающей - время" (рис. I) существует экстремальная линия, параллельная оси абсцисс, соответствующая минимуму болевого раздражения. Опыты показали, что огибающая ЭМГ в процессе опыта изменяется так (рис.I): она монотонно уменьшается до некоторого уровня, расположенного несколько ниже указанной экстремальной линии, достигнув этого уровня как бы "отскакивает" от него, чтобы вновь опускаясь, достичь этого уровня и вновь"отскочить" от него. Назовем такой "отскок" огибающей ЭМГ от указанного экстремального

уровня "всплеском". Опыты показывают, что эти всплески случайны по величине, но типичными являются две амплитуды всплесков - "малые" и "большие". Если в параболической характеристике, определяющей болевое раздражение, ликвидировать правую ветвь, то описанный процесс будет протекать точно также, но при этом исчезнут "большие" всплески и все всплески будут только "малыми". При неожиданном смещении экстремальной линии вниз или вверх (точка *а* и *б* на рис. 2) вся "картина" соответственно перемещается, и всплески происходят от новой линии, всегда расположенной ниже экстремальной линии и параллельно ей. В результате общая электрическая активность (и, соответственно, напряженность мышцы) постоянной колеблется вблизи значения, обеспечивающего минимум болевого раздражения, отслеживая его. Такое отслеживание происходит и при непрерывном перемещении экстремальной линии и ряд других фактов, связанных с этим процессом [1],

В случае, когда болевое раздражение зависит от большего числа мышц [2], процесс происходит аналогично, всплески этих мышц возникают одновременно. Прослежено, что всплески начинаются каждый раз, когда болевое раздражение, нарастая, достигнет некоторого порога, и не возникают, когда этот порог достигается при уменьшении болевого раздражения. Опыты показали также, что одновременно со всплесками рассматриваемых мышц возникают всплески и у иных "сторонних" мышц, от которых не зависит болевое раздражение, но точки, определяющие начало всплеска этих "сторонних" мышц, не регулярны, т.е. уже не ложатся на линии, соответствующие одинаковому значению болевого раздражения.

2°. Опыты по "прицеливанию" и предельно быстрому качанию сустава.

В этих опытах [3] записывалась ЭМГ и их огибающие для двух мышц-антагонистов какого-либо, например, лучезапястного, сустава и величина суставного угла. Эта величина наблюдалась в виде светящейся точки на осциллоскопе, а коэффициент усиления по углу выбирался столь большим, что изменение суставного угла на 2 - 3 угловых минуты уводило точку за пределы экрана осциллографа. При "точностном прицеливании" испытуемому ставилось за-

дание удержать точку в пределах экрана. При таком "точностном прицеливании" (рис. 3) суставной угол колеблется с частотой приблизительно 7 - 10гц. На огибающей ЭМГ обеих мышц четко наблюдаются типичные всплески, подобные тем, которые наблюдались в опытах с болевым раздражением (см. выше). Эти всплески при "точностном прицеливании" возникают каждый раз, когда абсолютная величина угловой скорости суставного угла, нарастая, достигает некоторого порогового значения. При этом всегда всплеск огибающей одной из мышц-антагонистов велик по амплитуде, а другой при этом - мал, или еле заметен. Большие всплески чередуются, т.е. если в некоторый момент был большой всплеск у сгибателя и малый у разгибателя, то в следующий раз, когда абсолютное значение угловой скорости достигает некоторого порога, наоборот, большой всплеск будет у разгибателя, а малый у сгибателя. Большой всплеск всегда возникает у той из мышц-антагонистов, которая в этот момент растягивается. При "точностном прицеливании" частота колебания суставного угла (7 - 10 гц)такова, что у каждой мышцы большой всплеск возникает примерно один раз за период, т.е. время между двумя всплесками колеблется от0,08 до 0,2 сек., типичное время - 0,1 сек.

Если теперь в условиях этого же опыта изменить задание испытуемому и потребовать, чтобы вместо точного прицеливания он качал руку в суставе с любой амплитудой, но предельно большой частотой, то запись окажется такой же, как и на рис. 3. Изменится лишь амплитуда колебаний сустава (а значит, слегка уменьшится частота) и уровень абсолютного значения угловой скорости суставного угла, при котором возникают всплески обеих мышц-антагонистов. Эта картина в общих чертах повторяется при опытах с разными суставами, при разных средних значениях суставных углов, около которых происходит качание при изменении в некоторых пределах нагрузки на кость, т.е. величины напряженности мышцы.

Сопоставление рис. 2 и 3 дает много оснований предполагать, что управление мышечной активностью в этих случаях происходит с помощью одинакового или однотипного механизма, что элементарной частью мышечной активности в этих случаях являются всплески, которые находят свое отражение во всплесках

огибающей ЭМГ, а управление ведется выбором моментов возникновения всплесков; эти моменты возникают каждый раз, когда некоторая величина, нарастая, достигает порога. Такой величиной в опытах п.I^0 являлось болевое раздражение, а в опытах п.2^0 - абсолютное значение угловой скорости сустава. Предлагаемые далее модельные представления объясняют, каким образом такое управление могло бы быть осуществлено. Авторы думают, что описанные выше опыты-лишь примеры работы рассматриваемого механизма.

§ 2. ПРОСТЕЙШИЙ ПОИСКОВЫЙ МЕХАНИЗМ (ППМ)

Центральными элементами рассматриваемого механизма являются случайный интернейронный пул (далее - СИП) и нервная организация, условно названная нами "функция-неприятность" (далее кратко - Н). Назначение элемента Н - вырабатывать в нужные моменты командные импульсные залпы. Назначение СИП - воспринимая залпы, отрабатывать то, что было названо выше всплеском.

I^0. Случайный интернейронный пул (СИП)

В качестве первичного параметра, характеризующего изменяющееся во времени состояние мышцы, мы рассматриваем мгновенное значение числа N_α возбужденных в этот момент α-мотонейронов. Несмотря на это, основную роль в предлагаемой модели будут играть не α- мотонейроны, а интернейронный пул мышцы. Именно, в качестве основной организации мотонейронного пула будет рассматриваться случайная сеть интернейронов (далее СИП), а α-мотонейроны будут пониматься лишь как "выходные реле" этой интернейронной сети.

Входами для СИП служат подходящие извне окончания, которые касаются случайного числа интернейронов данного СИП. При рассмотрении механизма ППМ существенную роль играют два входа в СИП: основной и фоновый. Эти входы отличаются не способом касания интернейронов (для обоих входов они случайны), а тем, откуда подводятся и какую роль играют проводимые по ним импульсы.

Л.И. Розоноэром [4] теоретически исследовались различные ре-

жимы работ СИП, причем в качестве "выхода" СИП считалось изменяющееся во времени число N_u возбужденных в каждый момент интернейронов. В силу того, что число α -мотонейронов достаточно велико, а подводы к ним от интернейронов случайны, числа N_u и N_α примерно пропорциональны.

Центральным для рассмотрения механизма ППМ является поведение СИП во время и после подвода по любому из входов короткого возбуждающего импульсного залпа.

Характер изменения N_u под действием одного импульсного залпа на входе может быть описан так: во время действия залпа N_u нарастает; по окончании действия залпа величина N_u уменьшается не мгновенно, а постепенно (и при том не обязательно монотонно), падая до того среднего значения, которое соответствует поведению автономного СИП и колеблется затем около этого значения (рис. 4).

Описанная реакция СИП на короткий импульсный залп называется далее "всплеском СИП". Наибольшее значение величины N_u, достигнутое во время всплеска, назовем величиной всплеска.

Из случайных свойств СИП непосредственно видно, что и величина, и длительность всплеска случайны.

Для работы ППМ всплеск является "атомом" мышечной активности, ее простейшим элементом. Работа ППМ целиком предопределяется тем, каким образом и в какие моменты посылаются импульсы для того, чтобы вызвать всплески, и мало зависит от длины залпа, числа импульсов в нем и их интенсивности. Если в качестве входного воздействия на СИП рассматривать лишь импульсацию, поданную к основному входу, то для такого воздействия подача дополнительного импульса на фоновый вход эквивалентна снижению среднего уровня порогов интернейронов СИП. Поэтому при одновременной подаче импульсов по фоновому входу число N_u будет существенно больше, чем было бы достигнуто при такой же импульсации на основном входе и отсутствии импульсации на фоновом входе. Благодаря этому, подачей сигнала на фоновый вход можно управлять величиной всплеска.

Все, о чем шла речь, связано лишь с работой той части механизма ППМ, которая обеспечивает возникновение случайных всплесков. Моменты же их возникновения предопределяются внеш-

ними для СИП импульсами и зависят от того, что выше было названо функцией "неприятность" - Н.

2°. "Функция неприятность" и условия возникновения командного импульса (закон всплеска).

Рассматривается лишь такая мышечная активность, в процессе которой мозг координирует и изменяет напряженность мышц так, чтобы было обеспечено поддержание некоторой наперед поставленной цели.

Когда отклонение от цели достигает некоторого "аварийного" порога, организм отвечает на это внешне наблюдаемой реакцией, например, каким-либо движением. Естественно поэтому считать, что существует "нечто", зависящее от рецепторных сигналов и нарастающее вместе с отклонением от цели, и что когда это "нечто" достигает "аварийного" порогового значения, возникают сигналы, вызывающие указанную выше наблюдаемую реакцию. Это "нечто" мы и называем "функция-неприятность" (или кратко - Н), но считаем, что, кроме указанного выше "аварийного" порога, существует у нее значительно более низкий порог, который используется механизмом ППМ для определения моментов выработки командных залпов (см. далее).

В этой работе мы не занимаемся локализацией нервной организации, вырабатывающей "функцию-неприятность", но учитываем, что в простейших случаях, когда "функция-неприятность" зависит от сигналов, поступающих только от рецепторов какого-либо одного типа (например, только от болевых рецепторов или от проприорецепторов мышцы или сустава), функция H может вырабатываться в самом рецепторе. "Функция-неприятность" определяет моменты выработки командных импульсов, которые направляются к спинальному сегменту мышцы или группы мышц (к основному входу СИП или, минуя СИП, непосредственно к α -мотонейронам).

Общий закон, в соответствии с которым определяются и изменяются эти моменты, мы назовем законом всплеска и сформулируем его так: импульсный залп возникает каждый раз, когда H, нарастая, достигает порогового значения $H_{пор.}$, и не возникает, если это пороговое значение достигается при уменьшении H. Значение $H_{пор.}$ само зависит от H_{min} уменьшаясь, если H_{min}

растет. В общем случае, когда нет информации о том, какие конкретные мышцы "повинны" в том, что H растет и достигает порогового значения, залпы направляются одновременно ко всем скелетным мышцам и вызывают у всех мышц одновременные, но случайные и разные по величине всплески. В тех же специальных случаях, когда есть информация о том, состоянием каких мышц определяется функция неприятность, команда о наступлении всплеска может направляться лишь к этим мышцам.

3°. Простейшая схема ППМ.

На рис. 5 показана простейшая схема ППМ. Подвод импульсного залпа к СИП вызывает всплеск СИП, т.е. всплеск числа N_α возбужденных α - мотонейронов. Всплеск СИП, т.е. нарастание и постепенное уменьшение числа N_α вызывает нарастание усилия, развиваемого мышцей, и постепенное ее расслабление до тех пор, пока подвод нового залпа от функции H на вход СИП (рис. 6) не вызовет нового всплеска N_α. В силу инерционности мышцы всплески ее напряженности более растянуты и сглажены по сравнению со всплесками СИП (рис. 6). Мышца по отношению к всплеску СИП является как бы сглаживающим инерционным фильтром.

Описанной простейшей схемы (рис. 5) достаточно, чтобы полностью объяснить экспериментальные результаты, полученные в опытах по поддержанию минимального болевого раздражения, когда оно зависит от напряженности мышц (см. § 2 п.I°). В этих опытах неприятностью служит само болевое раздражение. Начнем рассмотрение, считая, что в некоторый момент $t = t_0$ огибающая ЭМГ определяет точку a (рис. 7). В этот момент условий всплеска нет, число N_α (а значит, и огибающая) уменьшается до тех пор, пока не будет достигнуто некоторое значение y_1, при котором наступают условия всплеска (см.рис.7, точка a').

В этот момент подается залп (рис. 7) и возникает новый всплеск СИП. В процессе всплеска в соответствии с описанными выше законами работы СИП, нарастает, а затем уменьшается число N_α. Новый командный залп "сверху" наступит, когда при расслаблении мышцы вновь будет достигнуто значение y_1. Легко видеть, что таким образом полностью объясняется наличие малых и больших всплесков, переход на новый уровень за один всплеск,

тот факт, что всплески начинаются от линии, строго параллельной $y = y_{min}$ и всегда лежащей ниже ее, исчезновение "больших" всплесков при отсутствии правой ветви функции-неприятность и все остальные факты, обнаруженные в ходе этих опытов.

4°. Роль веретен и уточнение схемы ППМ. Работа схемы при управлении парой мышц-антагонистов.

Дальнейшее расширение возможностей ППМ происходит благодаря той особой роли, которую играет веретенная рецепция мышц, полностью игнорированная нами при предварительном рассмотрении механизма ППМ.

Напомним некоторые простейшие сведения о веретенной рецепции, которые далее нам будут нужны. Мышечные веретена являются рецепторами, реагирующими на: 1) изменение величины растяжения мышцы, т.е. на ее длину ℓ, и 2) одновременно на величины ℓ и $\frac{d\ell}{dt}$ (условно их их линейную комбинацию $h = a\ell + b\frac{d\ell}{dt}$). При этом коэффициент b заведомо нуль, когда $\frac{d\ell}{dt} \leq 0$, т.е. реагирование на скорость изменения длины происходит только при растяжении. Далее мы будем интересоваться лишь сигналом h.

Внешние командные воздействия на веретенный рецептор позволяют изменять коэффициенты a и b путем инервации специальных мышечных волокон (интрафузальных), расположенных внутри веретен, которые в общих чертах подобны обычным (экстрафузальным) мышечным волокнам. Напряженность их определяется в основном числом зажженных, непосредственно связанных с ними нейронов (в отличие от α-мотонейронов их называют γ-мотонейронами). Сигнал от веретен о величине h подводится как непосредственно к α-мотонейронам, так и к фоновому входу их СИП, о котором выше шла речь. Сигналы, управляющие числом зажженных γ-мотонейронов, подводятся "сверху", и мы предполагаем, в частности, что они могут подводиться также и от той нервной организации, где вырабатывается функция H. На рис. 8 показана схема механизма ППМ с учетом этих особенностей, вносимых веретенной рецепцией.

Представим себе теперь кость, поворачивающуюся в суставе благодаря действию только двух мышц-антагонистов (рис. 9).

Предположим сначала, что каждая из этих мышц управляется совершенно одинаковым механизмом ППМ. Общая для обеих мышц функция H зависит от величины или скорости, или одновременно от величины и скорости изменения суставного угла, и вырабатываемые в момент наступления всплесков импульсные залпы подаются одновременно к СИП обеих мышц. Предположим теперь, что в некоторый момент у функции H возникли условия всплеска. Командный залп подводится одновременно к СИП обеих мышц, но в результате возникают всплески разной величины, так как они зависят не только от командного залпа, но и от подпитки, поступающей по фоновому входу. Эта подпитка зависит от величины $h = a\ell + b\frac{d\ell}{dt}$ каждой из мышц-антагонистов порознь. Поэтому она у них не одинакова. Величина h в свою очередь зависит от того, в каком положении находится и куда движется в этот момент кость, т.е. от значений ℓ и $\frac{d\ell}{dt}$ у каждой из этих мышц.

Пусть, например, в момент подачи командного залпа h у левой мышцы на рис. 9 заметно больше, чем у правой мышцы. Тогда всплески обеих мышц произойдут одновременно, но величина всплеска у левой мышцы будет заметна больше, чем у правой, и следовательно, суммарное усилие, действующее на кость от обеих рассматриваемых мышц, будет направлено в сторону левой мышцы.

Эта неравномерность всплесков усиливается за счет упоминавшейся выше и нерассматриваемой пока реципрокной инервации (рис. 10).

Такая полная схема механизма ППМ для пары мышц-антагонистов полностью объясняет результаты упоминавшихся выше опытов (см.§2 п.2°) при точностном прицеливании и предельно быстром качании суставного угла. Неприятностью в этом случае служит абсолютное значение угловой скорости суставного угла. Поэтому командные залпы возникают каждый раз, когда абсолютная величина угловой скорости, нарастая, достигает некоторого порога, а подпитывается СИП той мышцы, которая в этот момент растягивается. Механизм полностью объясняет причину одновременного вспыхивания мышц-антагонистов, причину чередования больших и малых всплесков. Тремор кости является результатом этих всплесков: его основная, высокая (около 10 гц) частота предопределена частотой всплесков, а присутствующая в треморе медленная нерегулярная составляющая

результат того, что управление ведется по скорости, а величины всплесков случайны и нерегулярны. Описанный механизм полностью объясняет и иные детали, обнаруженные в ходе этих опытов (см.[3]). В частности, он объясняет полное совпадение наблюдаемых явлений при выполнении двух противоположных команд: предельный покой кости и, наоборот, ее предельно быстрое качание. В обоих случаях работает один и тот же механизм управления и при том совершенно одинаково†). В частности, с одной и той же функцией H – в обоих случаях ею служит абсолютное значение угловой скорости. С точки зрения этого механизма предельный покой есть просто быстрое качание, но только с малой амплитудой. Амплитуда зависит только от величины Δ функции H. Поэтому переход с одного режима на другой происходит просто изменением порога Δ у функции H, т.е. того уровня неприятности, при котором возникают командные импульсные залпы.

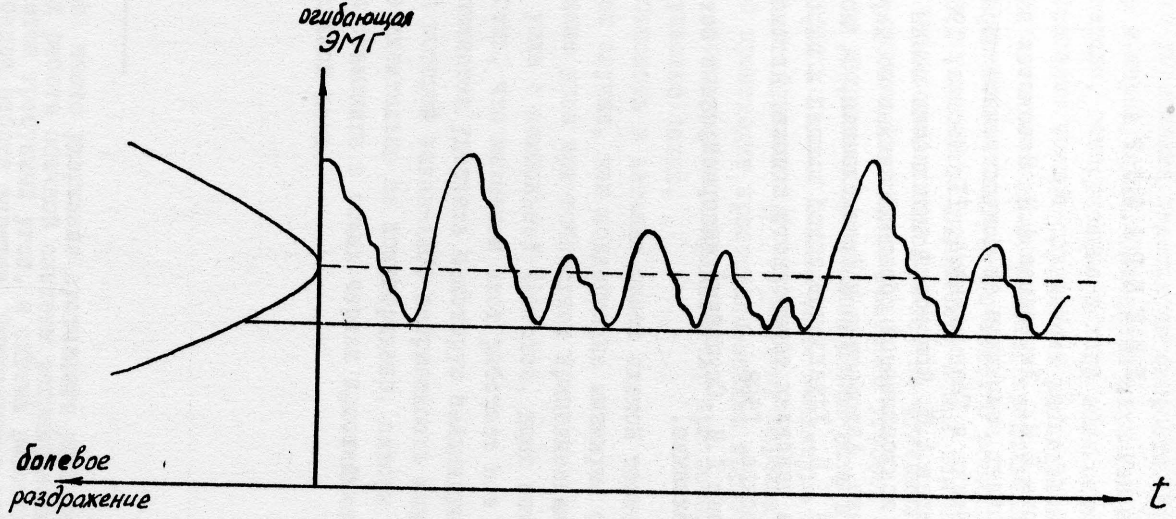
Возможность объяснить в деталях одной схемой такие три, казалось бы, различных случая, как поддержание минимума болевого раздражения, удержание угла при точностном прицеливании и поддержание колебаний угла с наибольшей частотой, дают авторам основание предполагать, что механизм этот работает и в иных случаях, когда целью является удержать некоторую найденную позу или найти позу, при которой какая-либо неприятность минимальна. Хотя это пока экспериментально не подтверждено, авторы считают, что механизм ППМ мог бы объяснить в общих чертах простейшие формы поддержания вертикальной позы.

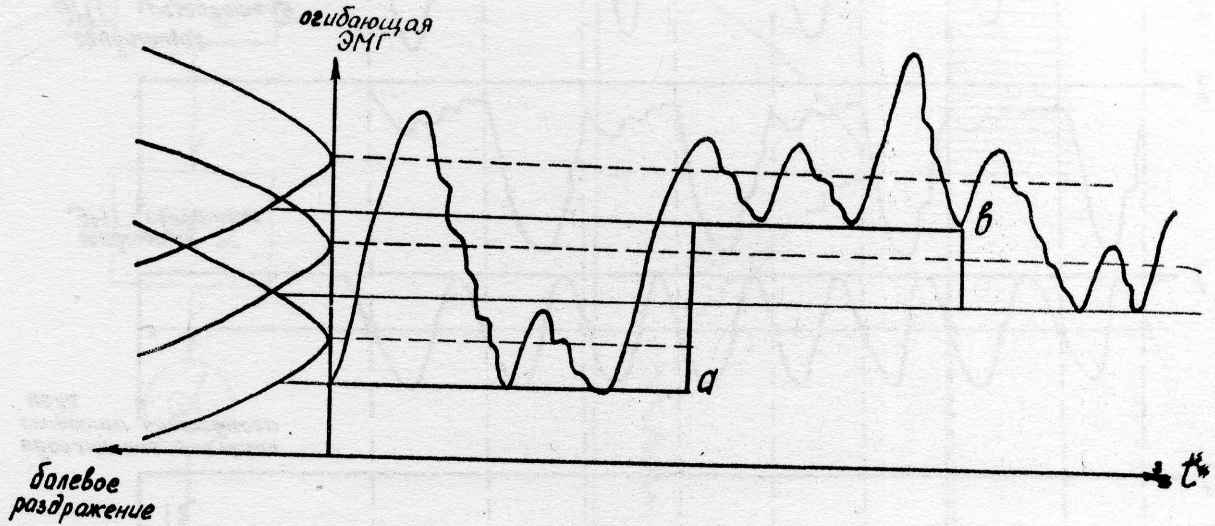
---

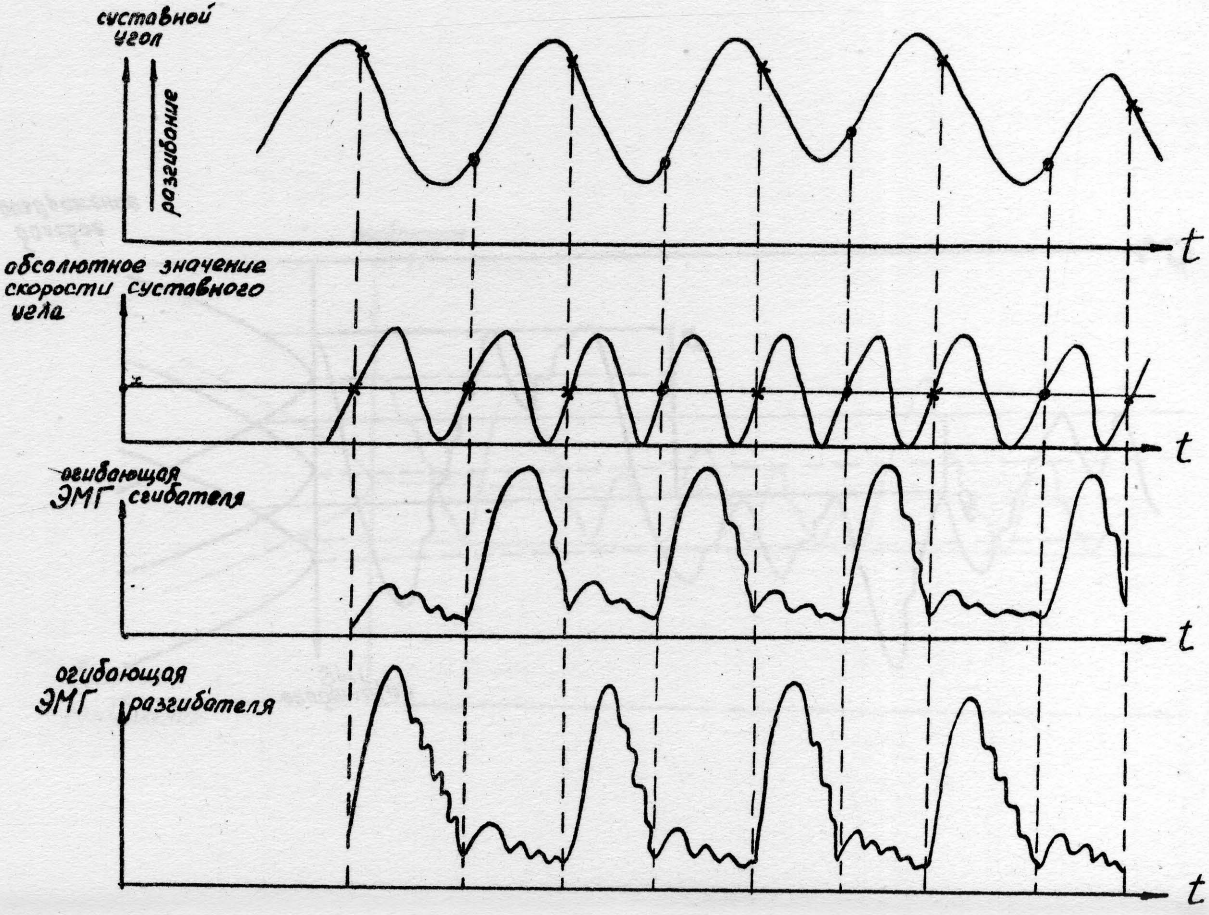
†) В литературе часто удержание суставного угла рассматривается как результат работы обычной системы регулирования, в которой "сверху" задается требуемый угол, а сигнал рассогласования получается благодаря работе веретен. Описанные в работе опыты опровергают такое предположение, т.к. оно не может объяснить одновременного вспыхивания мышц-антагонистов и то, что этот эффект особенно резко выражен при деафферентации мышц-антагонистов

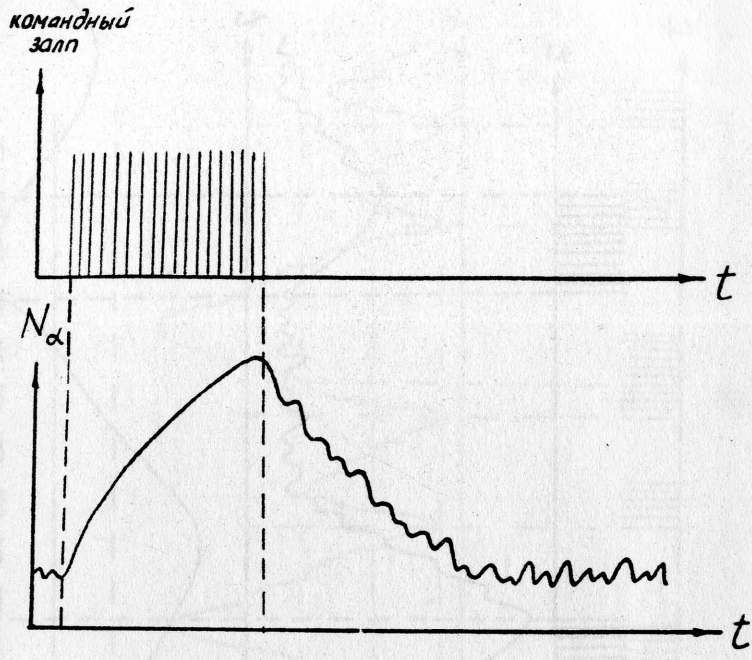
Л И Т Е Р А Т У Р А

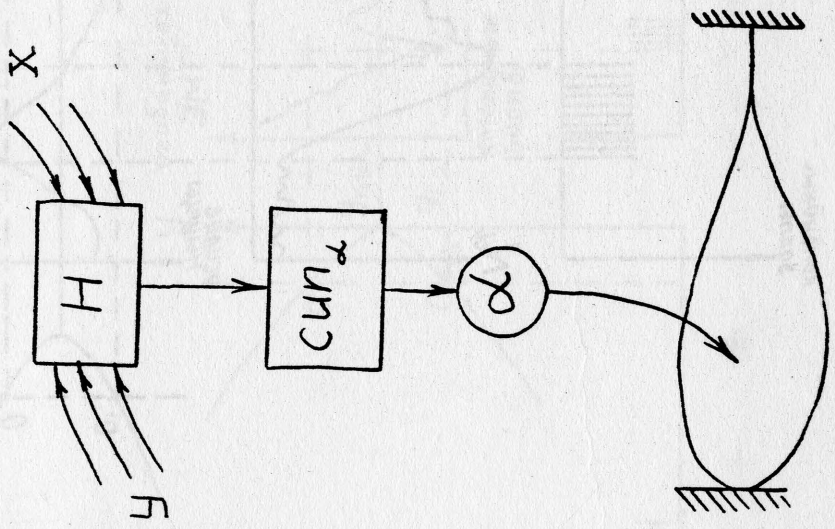
I. Захарова Л.М., Литвинцев А.И. Поисковая активность мышцы в условиях замыкания ее искусственной обратной связью, Автоматика и Телемеханика, II, Москва, I966.
2. Литвинцев А.И. Поисковая активность мышцы в условиях замыкания искусственной обратной связью по двум мышцам одновременно, Автоматика и Телемеханика, 3, Москва, I968.
3. Чернов В.И. Управление одной мышцей и парой мышц-антагонистов в условиях точностного поиска, Автоматика и Телемеханика, 7, Москва, I968.
4. Розоноэр Л.И. Случайные интернейронные сети, Автоматика и Телемеханика.

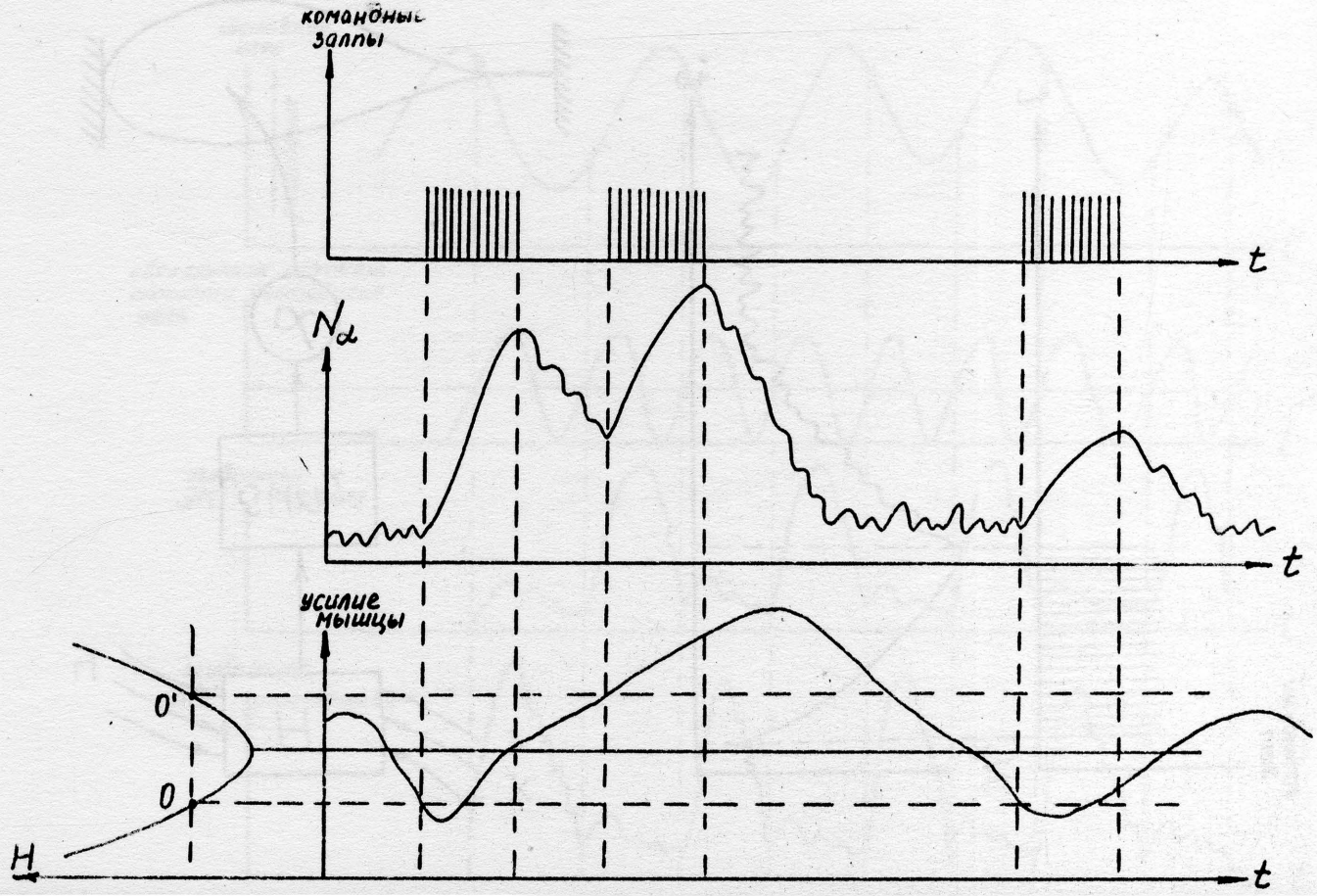

огибающая
ЭМГ
болевое
раздражение
t

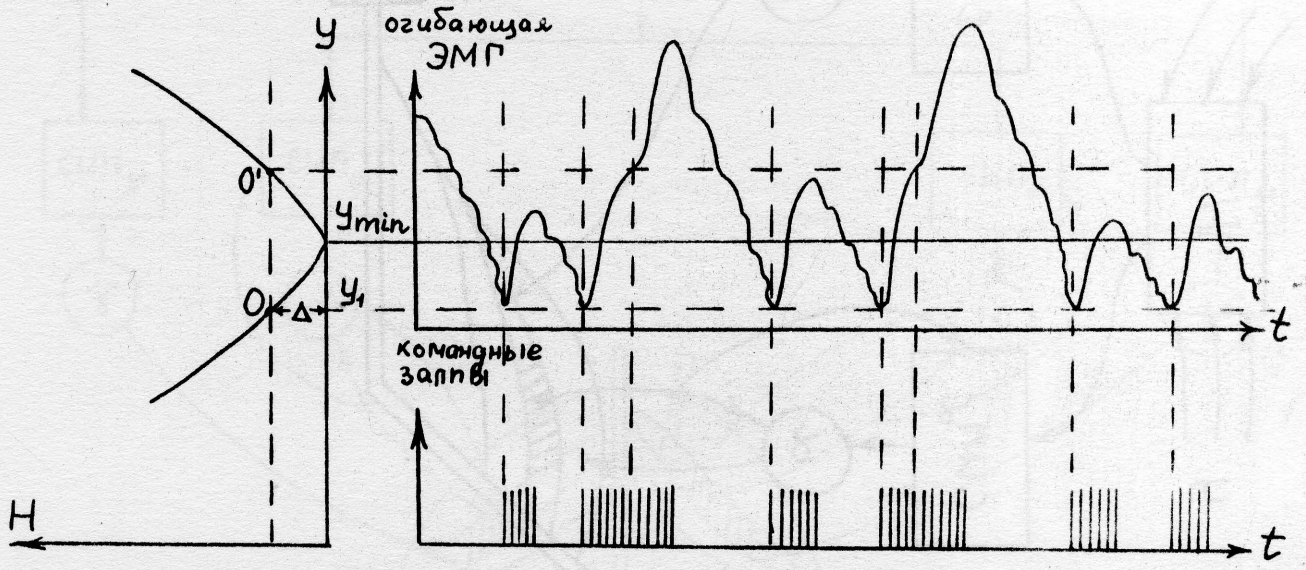

огибающая
ЭМГ
в
а
t
болевое
раздражение

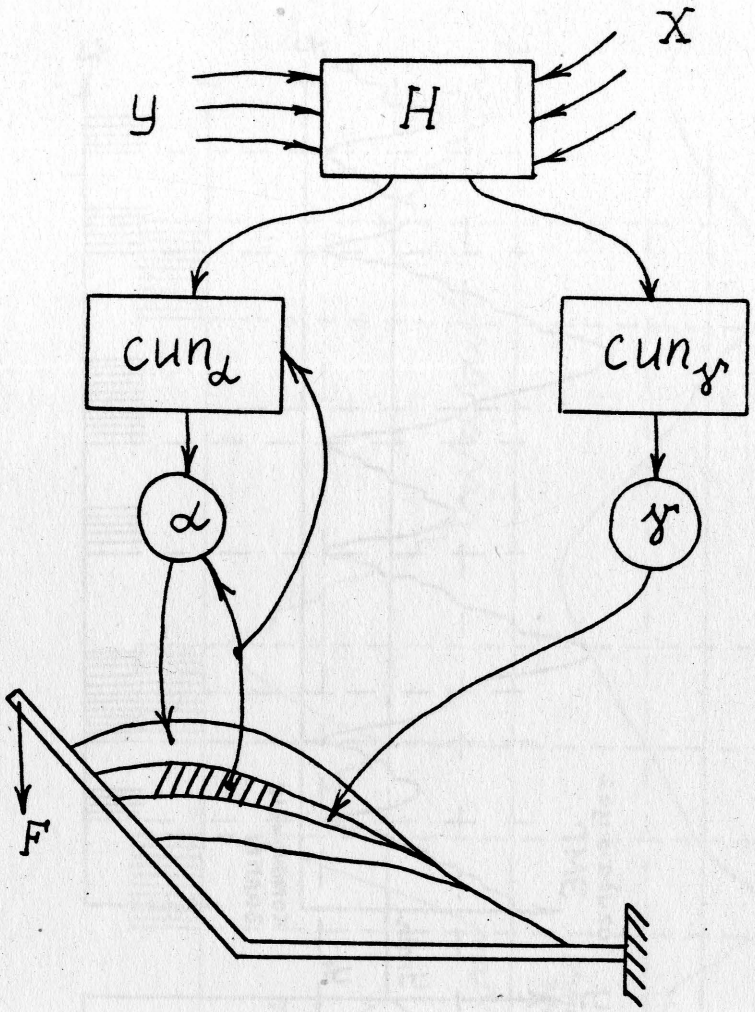

суставной угол
разгибание
t
обсолютное значение скорости суставного угла
t
огибающая ЭМГ сгибателя
t
огибающая ЭМГ разгибателя
t

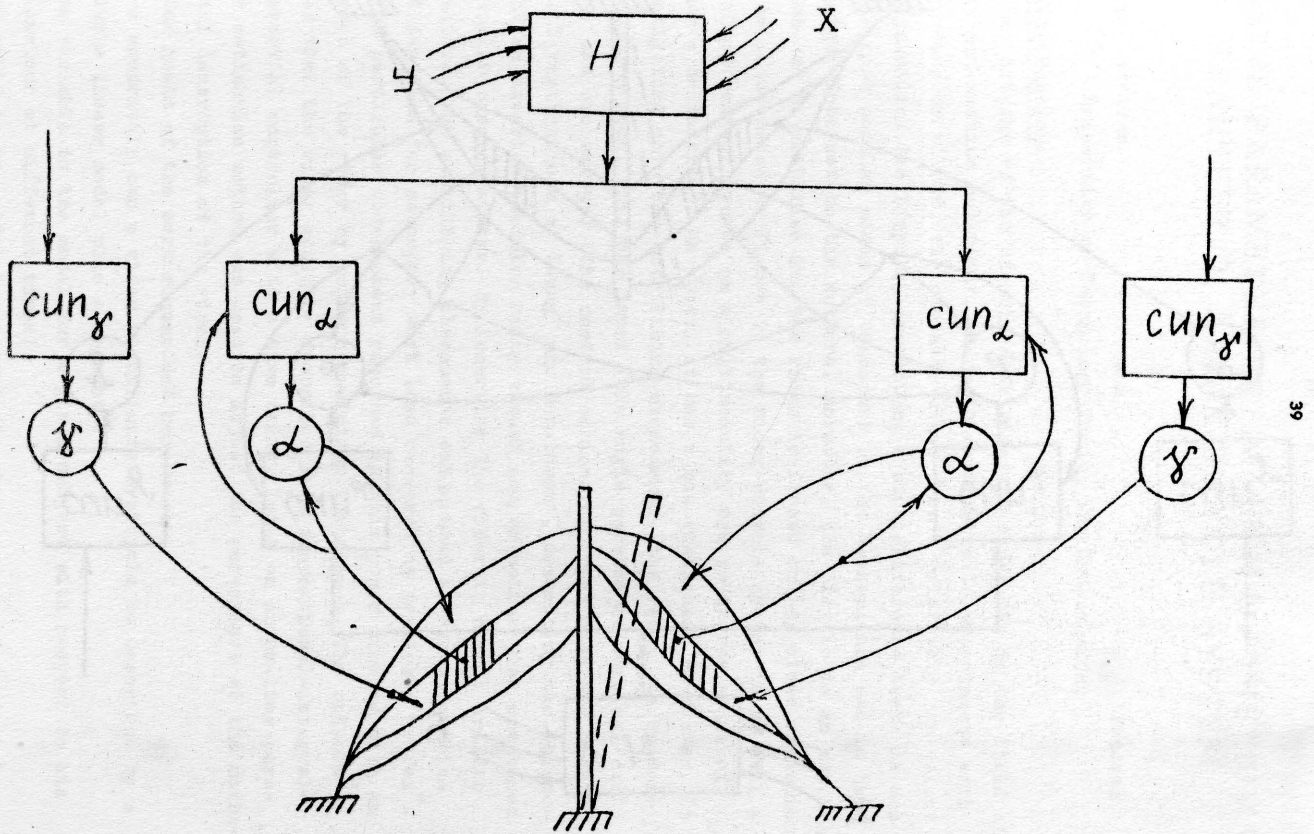

командный залп
t
N_α
t

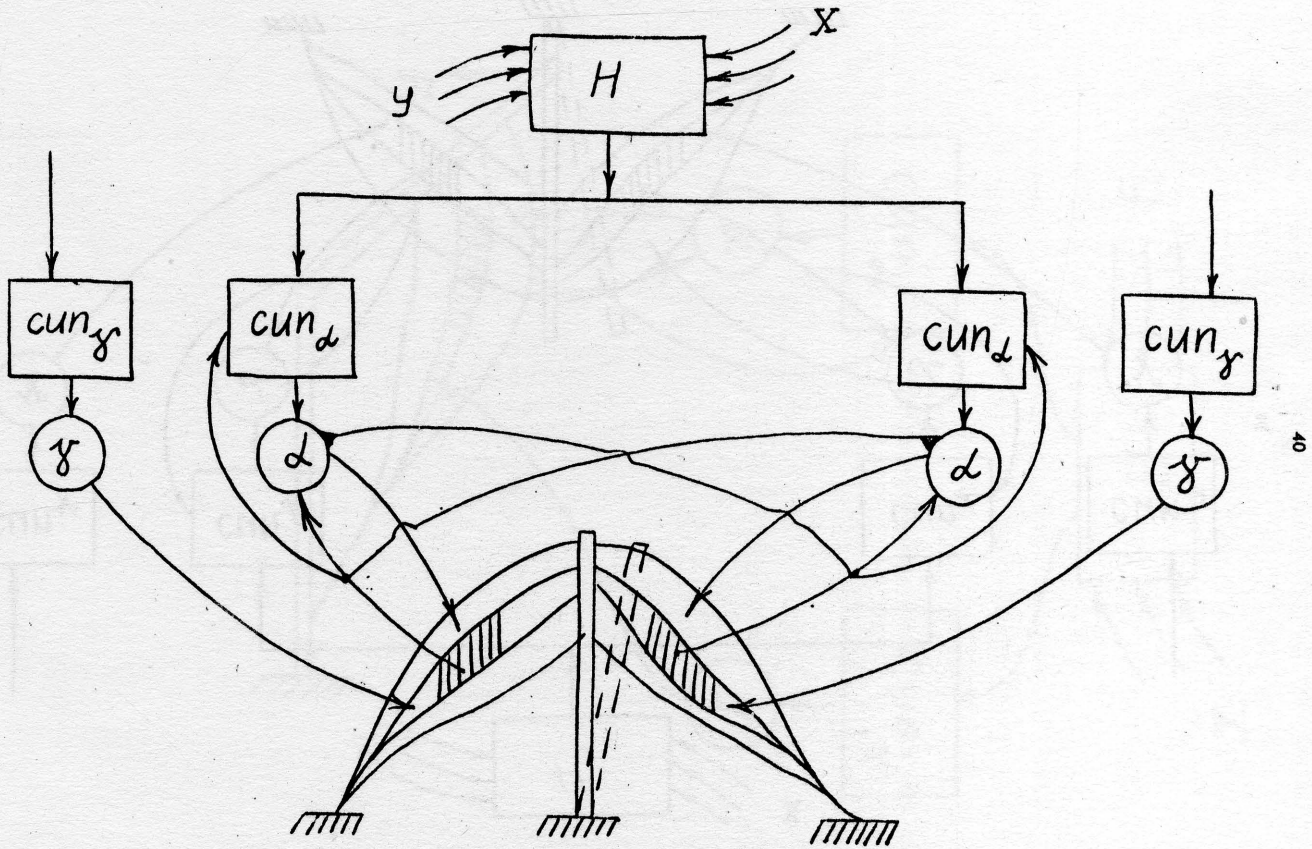

X
Н
$сил_\alpha$
α
Б


командные
залпы
t
N_α
t
усилие
мышцы
0'
0
H
t


y
огибающая
ЭМГ
0'
0
y_{min}
y_1
Δ
t
командные
запуы
H
t


X
y
H
сил$_\alpha$
сил$_\gamma$
α
γ
F


y
H
x
cun$_\gamma$
cun$_\alpha$
cun$_\alpha$
cun$_\gamma$
γ
α
α
γ


X
y
H
$сил_\gamma$
$сил_\alpha$
$сил_\alpha$
$сил_\gamma$
γ
α
α
γ

ON-LINE PARAMETER ESTIMATION OF THE HUMAN TRANSFER IN A MAN-BICYCLE SYSTEM

by

A. van Lunteren and H.G. Stassen

University of Technology Delft, The Netherlands.

1. Introduction.

The origins of the bicycle stabilisation study lie in the field of those man-machine relations, where in contrast with experiments with astronauts or pilots, the operator comes from a relatively non-selected population. The bicycle, presenting a straightforward task, is one of the most popular means of transport in the Netherlands and is thus of unique interest to the authors. Moreover, the bicycle is an unstable system, so to stabilise it, the rider is forced to pay close attention.

As illustrated in Figure 1, the rider-bicycle system is as well a multiloop control system as a multimodality system in the sense of McRuer's definition[1]. Moreover, it has a non-stationary character. In investigating only the stabilising phenomenon, the rider can be described by a model with one input (the angle between the vertical and frame) and two outputs (the control actions of upper body and handle bar). Investigations of Young[2] and Donaldson[3], describing the balancing of an inverted pendulum using visual and/or motion cues are based on similar considerations. By inserting the rider in normal traffic situations, the stabilising phenomenon can be used as a critical-instability task for secondary work load research as mentioned by Jex[4].

The particular development reported in this paper started with an analysis of the rider by using correlation techniques. The authors[5] showed that the rider could be described by a proportional-integral-differential controller with a time delay. Later on an on-line parameter estimation method was used to adjust the parameters of the mathematical description of the rider.

The goals of the work summarised here were:

a. How accurately can a rider stabilising a bicycle be described by a simple linear model plus a remnant[1]?

b. How variable is the model for a given subject with respect to his behaviour at different times?

c. How valid is the model for a randomly chosen group of subjects?

2. The bicycle simulator.

Based on a careful study of bicycle dynamics, published by Whipple

in 1899[6], a laboratory bicycle simulator was built. The following simplifications were assumed:

a. Only the stabilisation task on a straight course is considered.
b. The coupling between the rotation of frame and handle bar is neglected.
c. The body is considered to be consisting of two solid parts, viz. the upper and lower body.
d. Only small deviations about the equilibrium position are considered.

The rider-bicycle system is described by the linear differential Equation (1) derived from the moments acting on the frame about the ox axis through the contact points between wheels and ground (as modified from reference 6):

$$\underset{\text{GYROSCOPIC MOMENT}}{Av\dot{\alpha}_2(t)} + \underset{\text{CENTRIFUGAL MOMENT}}{Bv^2\alpha_2(t)} = \underset{\text{FRAME TILT}}{I\ddot{\varphi}(t)-Sg\varphi(t)} + \underset{\text{UPPER BODY}}{(I^*+zS^*)\ddot{\alpha}_1(t)-S^*g\alpha_1(t)} \qquad (1)$$

Figure 2 shows the axis systems used; the quantities are defined as follows:

$\alpha_1(t)$ is the angle between upper body and frame; $\alpha_2(t)$ is the angle between handle bar and frame; $\varphi(t)$ is the angle between the vertical oz axis and the frame; I and S are the moment of inertia and the static moment respectively of frame, wheels and the lower part of the human body about the ox axis; I* and S* are the moment of inertia and the static moment respectively of the upper body about an o*x* axis parallel to the ox axis and going through the saddle; z is the distance between the ox axis and the o*x* axis; A and B are constants, and v is the forward velocity.

Starting from the averaged values of weights and dimensions of the various parts of the human body[7] and considering a man with a weight of 63 kg and a height of 1.77 m., the numerical values for the constants of Equation (1) are: I = 109 kgm^2; Sg = 940 Nm; I^*+zS^* = 23.9 kgm^2; S^*g = 162 Nm; A = 25 kgm and B = 57.5 kg.

The laboratory bicycle simulator demonstrated a reasonable similarity to a normal bicycle; in general a rider was able to stabilise the simulator after a few minutes of practice. The forward motion was missing, but the effects thereof on the simulator were taken into account. The rotation of the shaft ox on which the frame was mounted, was accomplished by means of an electro-hydraulic servomotor. From the difference between the electrically generated gyroscopic and centrifugal moments, and the moment about the ox axis, measured by means

of a torsion transducer, the servomotor was controlled (see Figure 3). The characteristics of the bicycle model could be changed readily by varying the electronic components only. The pedaling torque could be generated by an electromotor, viz. in case of a motorised bicycle often called a moped; or by the rider, viz. in case of a normal bicycle. To measure human performance a test signal, simulating unpredictable sidewind variations, could be introduced into the system. Figure 4 shows a photograph of the simulator.

3. The mathematical model of the rider.

3.1. The mathematical model.

Human behaviour is strongly dependant on the system dynamics of the process to be controlled[8]. Investigations by Tustin[9] and Ragazzini[10] showed that the behaviour of a human operator in a man-machine system could be described by a linear model consisting of a PID-controller with a delay time, and in the context of the describing function method[11], an additive remnant. Other linear models, described by a transfer function with 2 zeros, 2 poles and a delay time, have been suggested[12,13,14]. The modelling technique mentioned in this paper is valid for a model with either poles or zeros only. A so called generalized model[15] implies an operation on the output as well as on the input of the human operator. However, in that case the parameter estimations are strongly biased by the remnant, even in the case of an open-loop system. Restricted by the computer available, only the PID-model could be applied to estimate the parameters on-line. In this manner the behaviour of a large group of subjects can be easily investigated. If the angle φ is considered as the input to the subject informing him about the state of the bicycle, and if the state variables α_j are considered as the control outputs of the subject ($j = 1$ for the motions of the upper body; $j = 2$ for those of the handle bar), then according to Figure 5 the form of the human transfer function will be:

$$H_j(s) = [C_{1j}(t)+C_{2j}(t)/s + C_{3j}(t).s]\ e^{-\tau_j(t)s} \qquad \text{----------(2)}$$

where s is the Laplace operator. $C_{1j}(t)$, $C_{2j}(t)$ and $C_{3j}(t)$ represent the proportional, integral and differential time varying constants respectively, and $\tau_j(t)$ the time varying delay times. To complete the describing function model of the bicycle-rider the remnants $\eta_j(t)$ are introduced.

3.2. The modelling technique.

The parameter estimation was achieved on-line by a digital computer as shown in Figure 5. Consider the digital model with input $\varphi[(m+k)\Delta t]$ and outputs $\alpha_j^*[(m+k)\Delta t]$ sampled at regular intervals Δt, where $m\Delta t$ is defined as the moment of first observation, and where the integer k is defined in the interval $o<k<n$. By defining the delay times $\tau_j = l_j\Delta t$, and according to Equations (2), the outputs of the model are given by:

$$\alpha_j^*[(m+k)\Delta t] = \sum_{i=1}^{3} C_{ij}[(m+k-l_j)\Delta t] \cdot \varphi_i[(m+k-l_j)\Delta t], \quad (3)$$

where $\varphi_1 = \varphi$, $\varphi_2 = \int \varphi dt$, and $\varphi_3 = \frac{d\varphi}{dt}$. The errors ε_j between the model outputs α_j^* and the rider outputs α_j are defined by Equation (4):

$$\varepsilon_j[(m+k)\Delta t] = \alpha_j[(m+k)\Delta t] - \alpha_j^*[(m+k)\Delta t]. \quad (4)$$

For selected arrays of delay times τ_j, assuming that the quantities C_{ij} are stationary during the time of observation $n\Delta t$, and using the quadratic performance criterion, the errors ε_j have to be minimised with respect to C_{ij} over the time $n\Delta t$, hence:

$$\frac{\partial}{\partial C_{ij}[(m+k-l_j)\Delta t]} \left[\sum_{k=o}^{n} \left\{ \varepsilon_j[(m+k)\Delta t] \right\}^2 \right] = 0. \quad (5)$$

By carrying out the differentiations of Equations (5), and taking into account Equations (3) and (4), there follows:

$$\sum_{k=o}^{n} \alpha_j[(m+k)\Delta t] \cdot \varphi_i[(m+k-l_j)\Delta t] + \\ - \sum_{r=1}^{3} C_{rj}[(m+k-l_j)\Delta t] \sum_{k=o}^{n} \varphi_i[(m+k-l_j)\Delta t]\, \varphi_r[(m+k-l_j)\Delta t] = 0. \quad (6)$$

For $l_j \ll n$ Equations (6) become:

$$\sum_{k=o}^{n} \alpha_j[(m+k)\Delta t]\, \varphi_i[(m+k-l_j)\Delta t] + \\ - \sum_{r=1}^{3} C_{rj}[(m+k-l_j)\Delta t] \sum_{k=o}^{n} \varphi_i[(m+k)\Delta t]\, \varphi_r[(m+k)\Delta t] = 0. \quad (7)$$

From this set of equations, and for a given array of l_j the coefficients C_{ij} can be calculated. Substituting Equation (3) into (4) and using $l_j \ll n$, the mean squared errors $\sigma_{\varepsilon_j}^2$, minimised to the coefficients C_{ij}, can be determined without knowledge of the actual values of ε_j, hence:

$$\sigma_{\varepsilon_j}^2 = \frac{1}{n} \left[\sum_{k=o}^{n} \{\alpha_j[(m+k)\Delta t]\}^2 - \sum_{i=1}^{3} C_{ij}[(m+k-l_j)\Delta t] \cdot \right. \\ \cdot \sum_{k=o}^{n} \alpha_j[(m+k)\Delta t] \varphi_i[(m+k-l_j)\Delta t] + \\ \left. + \sum_{i=1}^{3} \sum_{r=1}^{3} C_{ij}[(m+k-l_j)\Delta t] C_{rj}[(m+k-l_j)\Delta t] \cdot \varphi_i[(m+k)\Delta t] \cdot \varphi_r[(m+k)\Delta t] \right] \quad (8)$$

Thus, for the array of delay times, $0<l_j<L_j$ with $L_j \ll n$, the parameters C_{ij} with the corresponding mean squared errors $\sigma_{\varepsilon_j}^2$ can be found. Hence, the optimal delay times $\overline{l_j}\Delta t$ are now defined to be those delay times, belonging to the minimum values $\overline{\sigma_{\varepsilon_j}^2}$ of the quantities $\sigma_{\varepsilon_j}^2$. In this manner the optimal $\overline{C_{ij}}$ and $\overline{\tau_j}$ were found.

By dividing the observation time $n\Delta t$ into equal time spaces $r\Delta t$, in such a way that n/r is an integer and that $m = 0, r, 2r, 3r$, etc,

the coefficients $\overline{C_{ij}}$ and $\overline{\tau_j}$ can be determined over a time $n\Delta t$ at regular moments $m\Delta t$. Hence, the parameters are calculated by means of a running average procedure, so that the time varying character of $\overline{C_{ij}}$ and $\overline{\tau_j}$ can be investigated (see Figure 6). Finally four comments should be noted:

a. The uncertainty of the parameter estimation due to the finite time of observation decreases with a lengthening in this time. However, the uncertainty due to the non-stationary behaviour then increases. The optimal observation time has to be determined experimentally.

b. The sensitivity of C_{ij} to τ_j in the neighbourhood of $\overline{\tau_j}$ also has to be checked in practice, in this way getting information on the error in C_{ij} resulting from an error in τ_j.

c. The linear functions φ_i can be replaced by non-linear memory-less functions without any restriction.

d. The method is valid to determine the transfer function in an open loop system; in a closed loop system a bias due to the remnant is introduced.

3.3. Instrumentation.

The mathematical model was realised on a small digital P.D.P.-8 computer having a core-memory of 4096 words of 12 bits and a cycle time of 1.5 μsec., resulting in an addition time of 3.0 μsec. By means of a multiplexer the variables φ_i and α_j were fed into the computer. To remove drift and noise, the quantities φ and α_j were filtered by a first-order bandpass filter between 0.02 and 5 Hz. From the angle φ the variables φ_i were determined by analogue computation. The differentiating network had a break frequency at 50 Hz, while the integrating filter broke at 0.002 Hz. The sample frequency for each variable was 30 Hz.

The program of the computer generated the parameters C_{ij} for 5 values of τ_j as well as the variances σ of the variables φ, α_j and ε_j. The running averages of these quantities were determined every minute over $n\Delta t$ seconds. Manually the optimal values $\overline{\tau_j}$ were interpolated, while the optimal values for $\overline{C_{ij}}$ belonging to the optimal $\overline{\tau_j}$ were obtained from the quantities C_{ij}. To show the sensitivity of C_{ij} to τ_j a measure for the sensitivity, $\frac{\Delta C_{ij}}{30\ \Delta t}$, was calculated.
The test signal has been generated by a white noise source and has been filtered by means of a 4th order bandpass filter with a band width between 0.02 and 5 Hz.

4. Practical investigations.

4.1. The parameter estimation method.

To check the parameter estimation method described, the research was started by estimating the parameters of a simulated open-loop mathematical model between φ and α_1, which could be approximated by[5]:

$$\alpha_1(t) = \sum_{i=1}^{3} C_{i1}\,\varphi_i(t-\tau_1) + \eta_1(t). \qquad (9)$$

The transfer function was simulated on an analogue computer, as indicated in Figure 7. The variable φ and the remnant η_1 were obtained from gaussian white noise filtered by a fourth-order bandpass filter between 0.02 and 2 Hz. The spectra obtained in this manner were similar to those measured in bicycle experiments. The values of $\underline{C}_{11} = 0.25$, $\underline{C}_{21} = 0$, $\underline{C}_{31} = 0.10$ and $\underline{\tau}_1 = 0.1$ sec have been chosen according to an earlier result found with a bicycle experiment, viz.

$H_1(s) = -0.5(1 + 0.4s)\, e^{-0.1s}$.

A measure of relative remnant N_j; and relative error E_j is defined by Equations (10a) and (10b)

$$N_j = \frac{\sigma_{\eta_j}^2}{\sigma_{\alpha_j}^2} = \frac{\frac{1}{T}\int_0^T \eta_j^2(t)\,dt}{\frac{1}{T}\int_0^T \alpha_j^2(t)\,dt} \qquad \cdots (10a)$$

$$E_j = \frac{\sigma_{\varepsilon_j}^2}{\sigma_{\alpha_j}^2} = \frac{\frac{1}{n}\sum_{k=0}^{n}\{\varepsilon_j[(m+k)\Delta t]\}^2}{\frac{1}{n}\sum_{k=0}^{n}\{\alpha_j[(m+k)\Delta t]\}^2} \qquad \cdots (10b)$$

To show the influence of the magnitude of the remnant the level of N_1 was selected at $N_1 = 0.0$ and $N_1 = 0.4$. As indicated in section 3.3, the quantities $\bar{C}_{i1}$, $\bar{\tau}_1$, $\bar{E}_1$ and $\frac{\Delta C_{i1}}{30\Delta\tau_1}$ were calculated for a total observation time of 56 units of 1 min. each, therefore $n\Delta t = r\Delta t = 1$ min. Hence, for each quantity the mean values μ and the standard deviations σ_1 could be calculated. Moreover, the standard deviations σ_5 for the case where $n\Delta t = r\Delta t = 5$ min. were obtained. The results are given in Table 1. Figure 8 shows the mean squared relative error E_1 as a function of the delay time τ_1.

Taking into account that the time delay was approximated by a Padé-filter, Table 1 shows that for a stationary process in an open loop the parameters can be found to 10% accuracy even when the relative remnant N_1 is approximately 40%. Furthermore it is shown that an observation time of 5 min. was required for small standard deviations.

4.2. Analysis of the bicycle rider.

Only for zero remnants no difference will be found between open and closed loop systems. For non-zero remnants, application of this method in a closed loop implies that a bias due to the remnants is introduced. Experiments have suggested that for the ratios between remnants and side wind disturbance reported in this paper the estimation of the time delays have no significant deviations. Furthermore, the shape of the Bode plots shows only a slight difference between open and closed loop systems. Both the cases show that the integrating

action does not exist.

As shown for stationary processes, an observation time of 5 min. was satisfactory. To check whether the behaviour of the rider can be regarded as being stationary in this time interval, the parameters were determined for 2 subjects over 30 units of 1 min. each, i.e. $n\Delta t = r\Delta t = 1$ min. From the 30 values of each of the parameters the mean values μ and the standard deviations σ_1 were calculated. In the same manner as mentioned in section 4.1., the standard deviations σ_2, σ_3, σ_4 and σ_5 for 2, 3, 4 and 5 min. respectively were determined for 16 quantities $\bar{c}_{ij}$, $\bar{\tau}_j$, $\bar{E}_j$ and $\frac{\Delta c_{ij}}{30\Delta t}$. Furthermore, the standard deviations normalised to the one-min. case, σ_k/σ_1 for $k = 1, 2, 3, 4$ and 5 were calculated. In order to compress the results obtained, the mean value and the standard deviation of the normalised standard deviations σ_k/σ_1 were determined for each value k over the 16 quantities. By applying a similar procedure to the simulated rider of section 4.1., a comparison between the stationarity of the behaviour of the rider and the simulation could be made. The results shown in Figure 9 indicate that the behaviour of the rider can be regarded as stationary for an observation time of at least 5 min.

Finally, for a group of 5 male subjects, characterised in Table 2, the parameters were determined for a total observation time of 10 min. From the values of the first and last 5 min. the averaged values were obtained. The experiments were repeated each day for 10 days, between 2 p.m. and 3.30 p.m., under the same conditions. Each subject performed one experiment per day. In all experiments the simulator was used as a moped; v = 15 km/h. The mean values μ and the standard deviations σ_5 of the parameters over the 10 experiments for each subject are presented in Table 3. Moreover, the same quantities are given for the entire group in Table 4.

The results obtained show that the integration action for both handle bar and upper body can be neglected, hence the transfer functions based on the results of Table 4 become:

$$\left.\begin{aligned} \text{Upper body:}\quad H_1 &= -0.13\ [1 + 1.6\ s]\ e^{-0.09\ s} \\ \text{Handle bar:}\quad H_2 &= +1.07\ [1 + 0.15\ s]\ e^{-0.16\ s} \end{aligned}\right\} \qquad (11)$$

Taking the ratio between the differential constants $\bar{c}_{3j}$ and the proportional constants $\bar{c}_{1j}$ one observes that the lead time constant for upper body control is approximately 10 times larger than that for handle bar control. In contrast, the time delay for the handle bar control is about 1.7 times as big as that for the upper body control. Furthermore

one sees that the effective value of the remnant of the upper body control is about 50% of the total output, while that of the handle bar control is only 29%. Therefore, the linear fit for the handle bar action is much closer. Moreover, the standard deviations of the parameters for the proportional actions, in particular that of the constant $\overline{c}_{11}$, are much greater than those of the other parameters. The data suggest that the control of the upper body is more or less reflexive in nature, while the handle bar control involves higher levels of the CNS. Finally it can be seen that the standard deviation in the total group is greater than for an individual subject.

5. Concluding Remarks.

The parameter estimation technique is satisfactory in the sense that for a stationary proces in an open loop the parameters are generated with better than 10% accuracy for an observation time of 5 min. and a noise level of 40%. Although there is a slight difference in the case of a closed loop system, the authors believe that this length of time garantees enough freedom for future investigations of slow-time-varying phenomena, such as the influence of attention, fatigue and drugs.

The best linear fit for the rider was shown to be a proportional-differential controller with a time delay. Between the two control actions, significant differences were found for time delays, lead time constants, and effective values of the remnants. The results suggest that the control of the upper body is on a lower hierarchical level of the CNS than that of the handle bar. The high contribution of the remnants in the control of the bicycle justifies further study into the nature of this phenomenon. In particular, it appears worthwhile to investigate whether the remnant contribution is due to non-linear behaviour and/or to an introduced test frequency. Finally it should be mentioned that the standard deviations of the parameters of the total group were about twice as great as those of a single subject. This implies that a generalisation of the model to a group of subjects may not be justified.

6. Acknowledgements.

The authors would like to especially acknowledge Professor R.G. Boiten, who suggested the idea to study in detail the bicycle stabilisation phenomenon. Special thanks are also intended to A. J. Timmermans and F.J. Vergouwen for their invaluable contribution to the construction of the bicycle simulator.

7. References.

1. McRuer, D.T. et al.; New approaches to human-pilot/Vehicle dynamic analysis, Technical report AFFDL-TR-67-150 (1968), pp. 21.
2. Young, L.R. and Meiry, J.L.; Manual control of an unstable system with visual and motion cues, I.E.E.E. International Cong. Rec., vol. 13, pt. 6 (1965), p. 123 - 127.
3. Donaldson, P.E.K.; Error decorrelation studies on a human operator performing a balancing task, Med. Electr. and Biol. Engin., Vol. 2, nr. 4 (1964), p. 393 - 410.
4. Jex, H.R.; Two applications of a critical-instability task to secondary work load research, I.E.E.E. Transactions on Human Factors in Electronics, vol. HFE 8, nr. 4 (1967), p. 279-282.
5. Lunteren, A. van and Stassen, H.G.; Investigations on the characteristics of a human operator stabilising a bicycle model, Proc. Intern. Symp. on Ergonomics, Prague (1967), pp. 27.
6. Whipple, F.J.W.; The stability of the motion of a bicycle, Quarterly Journal of Pure and Applied Math., Vol. XXX (1899), p. 312 - 348.
7. Williams, M. and Lissner, H.R.; Biomechanics of human motion, W.B. Saunders CY, London (1962), p. 134 - 135.
8. McRuer, D.T. and Krendel, E.S.; A review and summary of Tracking research applied to the description of human dynamic response, Wescon Conv. Rec. (1958), p. 254 - 262.
9. Tustin, A; The nature of the operator's response in manual control and its implications for controller design, Journal of the I.E.E.E. vol. 94, pt. IIA (1947), p. 190 - 202.
10. Ragazzini, J.R.; Engineering aspects of the human being as a servomechanism. Unpublished paper, presented at the Am. Psychol. Assn. Meeting (1948).
11. Booton R.C.; The analysis of non-linear control systems with random inputs, Symp. on non-linear circuit analysis, Polytechn. Inst. of Brooklyn (1953), p. 369 - 391.
12. McRuer, D.T. and Graham, D.; Pilot-vehicle control system analysis, Progress in Astronautics and aeronautics, vol 13 (1964), p. 603 - 621.
13. Fogel, L.J.; An analysis for human flight control, I.R.E.-Convention Rec. pt. 8 (1956), p. 69 - 88.
14. Elkind, J.L.; Characteristics of simple manual control systems, M.I.T., Lincoln Lab. Tech. Report Nr. 111 (1956).

15. Eykhoff, P. et al.; Systems modelling and identification, proceedings Third IFAC Congress, survey paper, London (1966).

List of figures and tables.

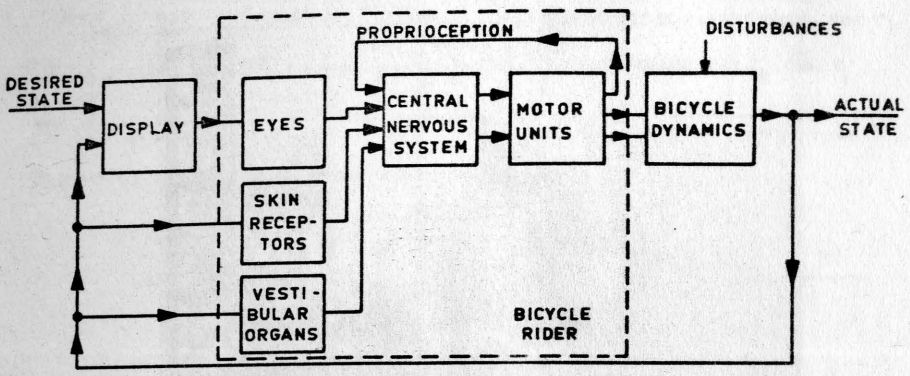


Figure 1: A general block diagram of the rider-bicycle system.

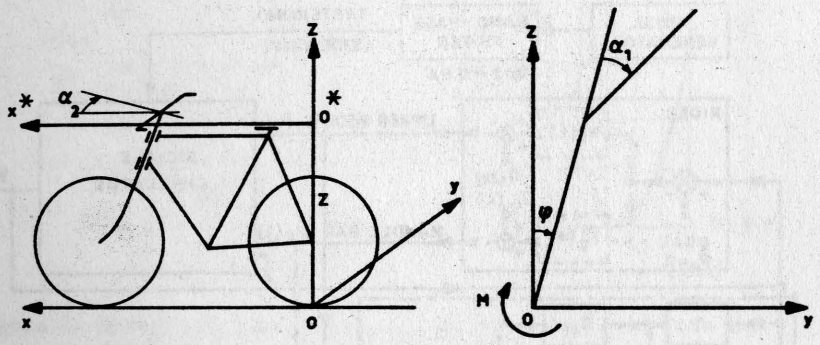


Figure 2: Definition of the variables of the bicycle simulator.

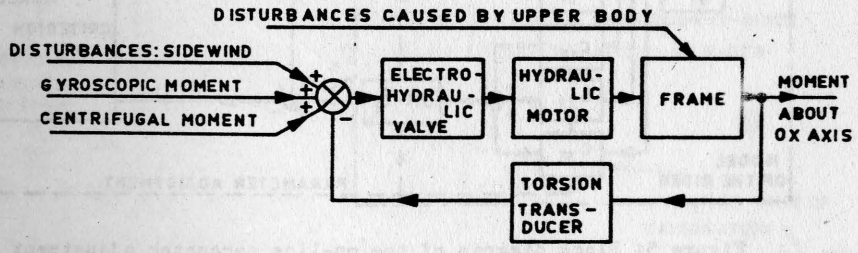


Figure 3: Block diagram of the control system of the bicycle simulator.

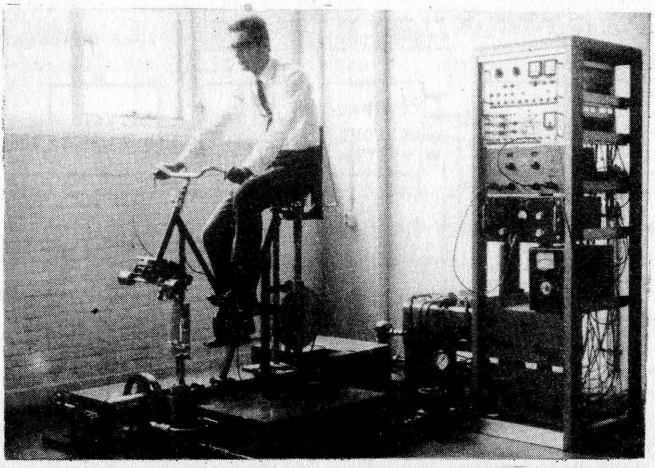


Figure 4: Photograph of the bicycle simulator.

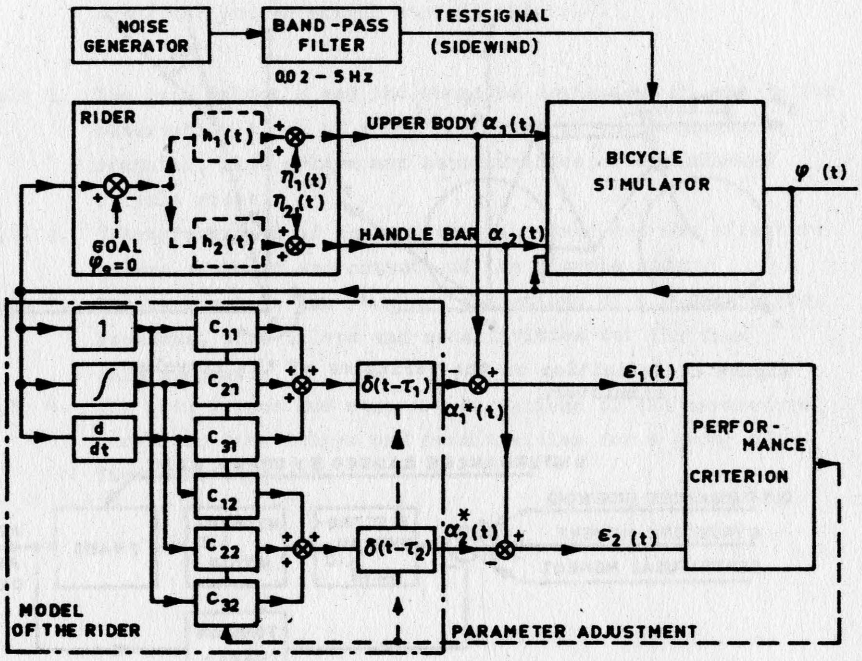


Figure 5: Block diagram of the on-line parameter adjustment of the mathematical model of the rider.

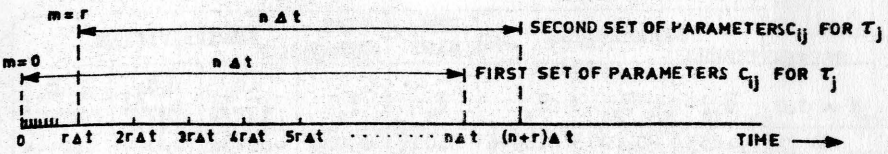


Figure 6: Timing of the parameter adjustment.

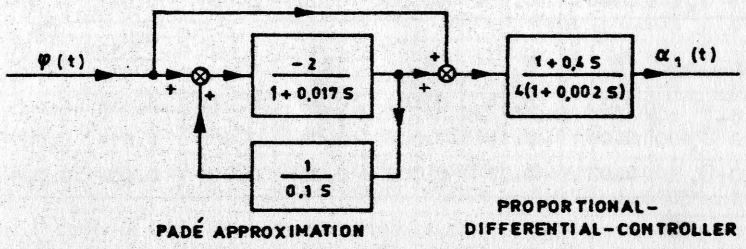


Figure 7: Simulation of the human transfer function of a bicycle rider.

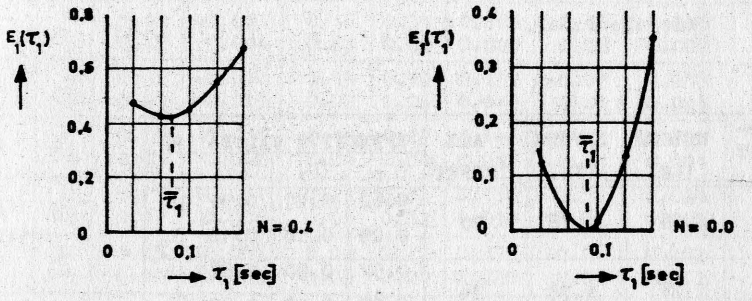


Figure 8: Example of the optimal values $\bar{\tau}_1$, calculated from the function $E_1(\tau_1)$ with $N_1 = 0.0$ and $N_1 = 0.4$.

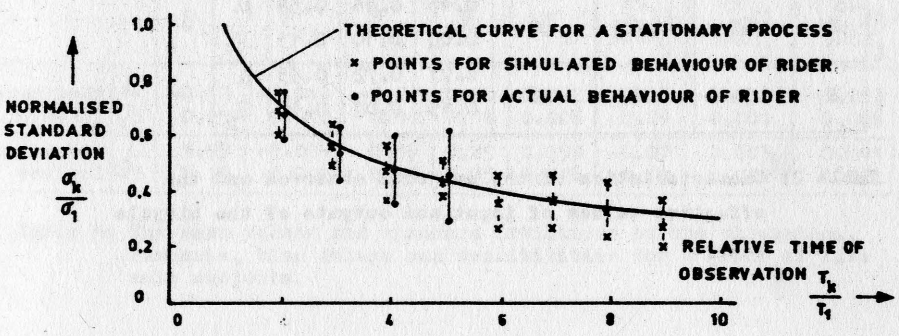


Figure 9: Comparison of the stationarity of the behaviour between a simulated and an actual bicycle rider.

	PARAMETERS			REM-NANT	TIME-DELAY	SENSITIVITIES		
N = 0.0	$\bar{c}_{11}$	$\bar{c}_{21}$	$\bar{c}_{31}$	$\bar{E}_1$	$\bar{\tau}_1$	$\frac{\overline{\Delta c_{11}}}{30\Delta\tau_1}$	$\frac{\overline{\Delta c_{21}}}{30\Delta\tau_1}$	$\frac{\overline{\Delta c_{31}}}{30\Delta\tau_1}$
μ	0.253	-0.007	0.0946	0.012	0.086	-0.186	0.004	0.0145
σ_1	0.002	0.004	0.0005	0.003	0.001	0.011	0.004	0.0005
σ_5	0.001	0.002	0.0003	0.001	0.001	0.005	0.001	0.0002
N = 0.4	$\bar{c}_{11}$	$\bar{c}_{21}$	$\bar{c}_{31}$	$\bar{E}_1$	$\bar{\tau}_1$	$\frac{\overline{\Delta c_{11}}}{30\Delta\tau_1}$	$\frac{\overline{\Delta c_{21}}}{30\Delta\tau_1}$	$\frac{\overline{\Delta c_{31}}}{30\Delta\tau_1}$
μ	0.23	0.14	0.100	0.38	0.084	-0.188	-0.003	0.011
σ_1	0.04	0.12	0.005	0.04	0.001	0.014	0.007	0.002
σ_5	0.02	0.07	0.002	0.02	0.001	0.006	0.004	0.001

Table 1: The mean values μ and the standard deviations σ_1 and σ_5 for observation times of 1 and 5 minutes of the parameters, remnants, time delays and sensitivities of a simulated bicycle rider.

SUB-JECT	WEIGHT [kg]	LENGTH [m]	AGE [years]	EFFECTIVE VALUES [o]			
				σ_φ	σ_α	σ_ψ	
A	66	1.78	20	0.62	0.91	0.97	μ
				0.09	0.18	0.10	σ_5
B	74	1.78	24	0.57	0.50	0.88	μ
				0.05	0.11	0.09	σ_5
C	63	1.63	23	0.75	0.57	0.72	μ
				0.11	0.10	0.06	σ_5
D	67	1.67	36	0.46	0.48	0.34	μ
				0.05	0.10	0.13	σ_5
E	71	1.83	26	0.71	0.72	0.95	μ
				0.15	0.07	0.10	σ_5

Table 2: Characteristics of the subjects observed and the effective values of input and outputs of the bicycle rider.

	SUB-JECT		PARAMETERS $\bar{c}_{11}$	$\bar{c}_{21}$	$\bar{c}_{31}$	REM-NANT $\bar{E}_1$	TIME DELAY $\bar{\tau}_1$	SENSITIVITIES $\frac{\overline{\Delta c}_{11}}{30\Delta\tau_1}$	$\frac{\overline{\Delta c}_{21}}{30\Delta\tau_1}$	$\frac{\overline{\Delta c}_{31}}{30\Delta\tau_1}$
upper body control	A	μ	-0.17	-0.004	-0.23	0.62	0.096	0.11	-0.003	-0.014
		σ_5	0.18	0.017	0.01	0.04	0.004	0.03	0.001	0.008
	B	μ	+0.17	-0.021	-0.25	0.48	0.095	0.11	-0.003	-0.010
		σ_5	0.09	0.007	0.02	0.04	0.003	0.03	0.001	0.005
	C	μ	-0.46	-0.005	-0.19	0.36	0.090	0.03	-0.001	-0.019
		σ_5	0.17	0.004	0.03	0.07	0.006	0.01	0.0004	0.005
	D	μ	-0.18	-0.020	-0.16	0.57	0.089	0.09	-0.002	-0.013
		σ_5	0.12	0.004	0.01	0.07	0.008	0.03	0.0005	0.004
	E	μ	-0.03	-0.020	-0.24	0.48	0.091	0.08	-0.002	-0.010
		σ_5	0.12	0.010	0.02	0.05	0.004	0.04	0.001	0.005
	SUB-JECT		$\bar{c}_{12}$	$\bar{c}_{22}$	$\bar{c}_{32}$	$\bar{E}_2$	$\bar{\tau}_2$	$\frac{\overline{\Delta c}_{12}}{30\Delta\tau_2}$	$\frac{\overline{\Delta c}_{22}}{30\Delta\tau_2}$	$\frac{\overline{\Delta c}_{32}}{30\Delta\tau_2}$
handle bar control	A	μ	1.12	-0.003	0.17	0.31	0.161	-0.08	0.003	0.047
		σ_5	0.12	0.004	0.03	0.03	0.009	0.02	0.002	0.012
	B	μ	1.17	-0.008	0.14	0.30	0.159	-0.07	0.003	0.044
		σ_5	0.09	0.003	0.02	0.06	0.014	0.03	0.003	0.008
	C	μ	0.78	-0.000	0.15	0.22	0.146	-0.03	0.001	0.029
		σ_5	0.03	0.002	0.04	0.04	0.012	0.02	0.001	0.006
	D	μ	1.23	-0.001	0.21	0.32	0.156	-0.12	0.004	0.045
		σ_5	0.09	0.008	0.04	0.02	0.010	0.05	0.002	0.014
	E	μ	1.05	-0.006	0.14	0.29	0.171	-0.06	0.002	0.040
		σ_5	0.14	0.003	0.03	0.04	0.020	0.03	0.002	0.007

Table 3: The mean values and standard deviations of the parameters, remnants, time delays and sensitivities for the five male subjects.

		$\bar{c}_{1j}$	$\bar{c}_{2j}$	$\bar{c}_{3j}$	$\bar{E}_j$	$\bar{\tau}_j$	$\frac{\overline{\Delta c}_{1j}}{30\Delta\tau_j}$	$\frac{\overline{\Delta c}_{2j}}{30\Delta\tau_j}$	$\frac{\overline{\Delta c}_{3j}}{30\Delta\tau_j}$
upper body control	μ	-0.13	-0.014	-0.21	0.50	0.092	0.08	-0.002	-0.013
	σ_5	0.25	0.012	0.04	0.10	0.006	0.04	0.001	0.006
handle bar control	μ	1.07	-0.004	0.16	0.29	0.159	-0.07	0.003	0.041
	σ_5	0.19	0.005	0.05	0.05	0.016	0.04	0.002	0.012

Table 4: The mean values and standard deviations of the parameters, remnants, time delays and sensitivities for a group of five male subjects.

DYNAMICS OF DRIVER/VEHICLE STEERING CONTROL

David H. Weir, Principal Research Engineer
Duane T. McRuer, Technical Director
Systems Technology, Inc., Hawthorne, Calif., U. S. A.

INTRODUCTION

Some form of driver steering control is required to maneuver a motor vehicle or to regulate its path. To understand driver control quantitatively one must consider the actions of random and deterministic inputs on a dynamic system comprising a vehicle whose equations of motion are known and a driver whose response under various conditions can be estimated. This paper describes some features of driver guidance and control, and summarizes the derivation of operational models for the driver/vehicle/roadway system.[1,2]

The general topology of the closed-loop structure for steering control is presented in Fig. 1. This depicts the possible types of driver response blocks, how the driver interacts with the vehicle, and how the driver/vehicle system interacts with the roadway environment. The controlled element contains the dynamics of the vehicle and the steering system, as well as the geometry of the visual field from which the driver must extract the guidance and control cues. The roadway environment provides inputs to the system, including both commands to be followed and disturbances to be regulated against. All the blocks in Fig. 1 contain a frequency or time function for modelling purposes. In fact, the essence of this approach is the derivation and exercise of these functions and their interaction during various driving situations.

The development of driver/vehicle dynamic models is based partly on a prior understanding of the manual control of aircraft. The fields have many common features, and pertinent results from flight simulators have now been applied to driving. The vehicle equations of motion have a similar form, but differ in detail. The pilot's perceptual field during visual (VFR) approach and landing is analogous to that in driving, yet each task has its own complexities. Night driving involves some visual field abstraction, while instrument flight and fire control tasks use artificial displays of needed cues. In each case good closed-loop system characteristics are essential, and the human operator adapts his response and behavior to suit the vehicle dynamics and task variables.

---

*This paper describes research results which are derived in part from work accomplished under subcontract to Contract No. CPR-11-2770, "Defining Requirements of Overtaking and Passing Maneuvers," with the Bureau of Public Roads. The prime contractor was The Franklin Institute Research Laboratories.

VEHICLE DYNAMICS

The vehicle dynamics are an important part of the effective controlled element, Y_c. The driver/vehicle/roadway model examined here considers the directional or steering dynamics in three degrees of freedom, i.e.,

- v lateral (or side) velocity
- r heading angle rate
- φ sprung mass roll angle

Other motions of interest (see Fig. 2) include:

- ψ heading angle
- γ path angle
- y_I inertial lateral deviation or position (of c.m.)
- a_y lateral acceleration (at various locations)

Another useful motion quantity is $y_I(t+T)$, the vehicle's lateral position T seconds ahead if it continues unperturbed along its present trajectory.

Vehicle motion-to-steer-angle (δ_w) transfer functions have been computed[1] for a typical medium-sized American sedan (circa 1965) weighing about 4,000 lb at a speed of 60 mph. The denominator is

$$\Delta(s) = .684[s^2 + 2(.785)(5.84)s + (5.84)^2][s^2 + 2(.290)(6.97)s + 6.97)^2] \quad (1)$$

The first mode is associated primarily with heading and sideslip while the second is dominated by the roll degree of freedom. Numerators for lateral velocity, heading rate, and lateral acceleration at c.g. are

$$N^v_{\delta_w} = 180(s - 9.50)[s^2 + 2(.284)(6.94)s + (6.94)^2] \quad (2)$$

$$N^r_{\delta_w} = 24.3(s + 4.35)[s^2 + 2(.284)(6.95)s + (6.95)^2] \quad (3)$$

$$N^{a_y}_{\delta_w} = 180[s^2 + 2(.293)(6.91)s + (6.91)^2][s^2 + 2(.160)(7.24)s + (7.24)^2] \quad (4)$$

Path angle is the first integral of lateral acceleration at the center of mass, and inertial lateral deviation is the double integral. Heading angle is the integral of heading rate. There is little difference between the dynamics at 60 and 75 mph, indicating that one operating point can be used to analyze typical high speed driving maneuvers which involve speed changes of no more than about 15 – 20 mph.

The series dynamics of the steering system are an important part of the effective controlled element. Although the theory is well understood,[3] dynamic data are sparse, and a pure gain is used in this discussion.

DRIVER DYNAMIC RESPONSE

All phases of driving require some form of driver control operation. As the tasks become more demanding, the driver may change his dynamic characteristics or may alter the system structure (close other loops) to obtain the required increase in control fidelity. The driver's closure of feedback loops modifies the effective dynamics of the vehicle or controlled element and in turn determines maneuver times, stability margins, and transient response characteristics. The possible feedback loops he can introduce are determined by the sensory information available.

Several possible types of driver response are presented in Fig. 1, corresponding to skill levels in the "Successive Organization of Perception,"[4] i.e.,

> **Compensatory** which implies an operation on a perceived error between the actual vehicle motion and the desired motion or input quantity.
>
> **Pursuit** which takes advantage of a knowledge of the system input to structure a driver feedforward which improves performance.[5]
>
> **Precognitive** which involves executing a learned maneuver in an open-loop way.

The first requires a fairly complex description, but it is well understood at the current time and probably comprises a significant portion of the driver's active control efforts. The remaining two are somewhat easier to describe qualitatively, but predictive models for these processes are not yet well developed. The quasi-linear blocks in the driver/vehicle model are most appropriate for response to random-appearing external inputs, including commands due to bends and curves as well as disturbances due to gusts and roadway roughness. These blocks are relatively quiescent in the presence of deterministic inputs when the other types of response dominate. They are active during pursuit control if random-appearing disturbances which the driver cannot preview are present, also.

Quasi-Linear Compensatory Control

The quasi-linear describing function model of the operator has resulted from an exhaustive series of human operator dynamic response measurements made over a period of about two decades.[6-9] It consists of a describing function component with parameters which depend on the system and situation, a set of rules which tell how to adjust the parameters, and an

additive remnant. This quasi-linear model is shown in the single-loop block diagram of Fig. 3. In its most complete form the describing function contains gain, indifference threshold, time delay, equalizer, and neuromuscular dynamics. The indifference threshold is a higher order effect that can often be ignored when the inputs are large and under other conditions it can be accounted for by using decreased driver gain. The neuromuscular system dynamics are based on very low and very high frequency data, and can be approximated at the mid-frequencies of interest in driving as a first-order lag or even as an added increment to the time delay. With these simplifications, the general driver describing function reduces to:

$$Y_p \doteq K_p \left(\frac{T_L j\omega + 1}{T_I j\omega + 1}\right) e^{-j\omega(\tau + T_N)} \qquad (5)$$

where K_p is the gain

$\left(\frac{T_L j\omega + 1}{T_I j\omega + 1}\right)$ is a simplified equalization characteristic

τ is the time delay

T_N is the neuromuscular system time constant

The form of Eq. 5 is adequate for a variety of drivers, inputs, vehicle dynamics, steering system characteristics, and loop structures. Most of the parameters are adjustable as needed, and typical values are illustrated by the subsequent system analyses.

The equalizing characteristic and the gain, K_p, are the major adaptive elements of the human which allow him to control many differing dynamic devices. The major "adjustment rules" are that a particular equalization is selected from the general form $K_p(T_L j\omega + 1)/(T_I j\omega + 1)$ such that:

- The driver/vehicle system can be stabilized
- The amplitude ratio of the product of driver and vehicle, $|Y_p Y_c|$, has approximately a −20 dB/decade slope in the crossover region
- $|Y_p Y_c| \gg 1$ at low frequencies

The major "cost" of equalization is an increase in effective time delay incurred when low frequency lead is needed as part of the compensation.

Simplified Crossover Model for the Driver Describing Function. The rationale of driver equalization can be adequately and simply expressed by using an approximate "crossover model."[6] Both experimental data and consideration of the requirements of good feedback system performance lead to the conclusion that the driver adjusts his describing function, Y_p, such that the open-loop function, Y_pY_c, in the vicinity of the gain crossover frequency, ω_c, has the approximate form

$$Y_pY_c \doteq \frac{\omega_c e^{-j\omega\tau_e}}{j\omega} \qquad (6)$$

where τ_e is an effective pure time delay which includes the neuromuscular time constant, T_N, as well as τ and any net high frequency controlled element lag. The gain term is the crossover frequency. The driver adopts either proportional control, or lead or lag equalization such that the product of the equalization and the vehicle has the form shown, with which it operates on the perceived motion error. The numbers (ω_c and τ_e) in the crossover model depend on the driver equalization. The crossover frequency is greatest and the effective time delay is least when the driver's equalization is a low frequency lag. Alternatively, ω_c is least and τ_e the greatest for low frequency lead. Thus the closed-loop bandwidth and performance will be reduced for low frequency lead and greatest when only lag is needed. Experimental values for the crossover model parameters for several dynamic forms and input bandwidths are given in Ref. 6. These are maximum values for skilled subjects in fixed-base simulators without motion feedbacks, and will ordinarily be considerably reduced during driving. This leads, in practice, to larger phase and gain margin criteria, typical examples of which are used in the subsequent driver/vehicle loop closures.

Driver Remnant. The remnant is that portion of the driver's output which is not linearly correlated with the input. It is considered to be an additional random process, denoted by a power spectral density, Φ_{nn}, in Fig. 3. Its major source appears to be nonstationarity in the operator's behavior.[6] For nominally good vehicle dynamics the remnant power is about 30 dB down relative to the input, and can usually be neglected from the standpoint of predicting driver/vehicle closed-loop response characteristics.

Driver Response in Multiple-Loop Situations. The system structure of Fig. 1 is multiloop, involving angle and path feedbacks. Experimental

measurements[9] of operator response in such a multiloop situation show that his describing function in the outer loop can be obtained by application of the single-loop model, above. The inner loops supplied by the driver act as equalization for the outer loop, or provide feedbacks or crossfeeds which suppress subsidiary, undesirable controlled element degrees of freedom. Because their role is so dependent on outer-loop requirements, the single-loop model rules are not generally applicable. Data indicates that the types of inner loops closed and their equalization are compatible with:

- Making outer-loop adjustments per the single-loop adjustment rules more feasible.
- Reducing the sensitivity of the closed-loop system to changes in either inner- or outer-loop driver response
- Selecting a loop structure and equalization which gives best subjective opinion rating[10]

It frequently happens that the driver describing function synthesized for the inner loop alone via the single-loop model is also the one which best enhances the outer-loop closures; particularly with a good stable basic vehicle, and a control task which is merely following command inputs or suppressing disturbances.

Pursuit Control and Internally Generated Maneuvers

The pursuit control block in Fig. 1 operates on the input using the driver's preview of a desired path, rather than a perceived motion or path error as in the compensatory case. Also, for some highly skilled maneuvers the driver produces the appropriate steering action based on an internally generated pattern previously evolved during a learning process. Recent operator response experiments[5] and analyses[11] have concentrated on deriving models for such systems. Although some useful results have been obtained, their application to the prediction of driver behavior is not yet state-of-the-art. However, a compensatory loop is active even when pursuit behavior is exhibited, so the compensatory quasi-linear model can be used to provide a first-order approximation to many situations. This will be the procedure adopted here.

STRUCTURES OF DRIVER/VEHICLE CLOSED-LOOP SYSTEMS

Perhaps the most difficult problem in driver/vehicle closed-loop analysis is to determine what sensory feedback loops the driver is using. There are two general approaches to deriving the closed-loop structure:

from experimental perceptual studies and from guidance and control theory. The theoretical approach considers the kinds of vehicle motions which must be sensed and commands which must be inserted to satisfy the guidance and control needs. This does not lead to a unique solution, so the initial results are a number of "sufficient systems." Consideration must then be given to the driver's ability to perceive (sense) the vehicle motion and input quantities, and to the nature of the operations required on the sensed quantities in closing the driver/vehicle system loops. Those potential systems which involve readily sensed quantities and good system performance are then accepted as suitable candidates. The control theory approach is relatively straightforward, given the current state of knowledge of typical operator/vehicle control laws and closed-loop analysis techniques. The best approach is to marry the guidance and control and the perceptual theories, for the net results must be compatible.

Single-Loop Structure

The problem of how the driver perceives a candidate feedback variable is secondary in the deductive guidance and control approach. The objective is to discover likely good loops and determine how well the motion variable considered permits the driver to control the vehicle if it can be sensed. It is also desired to identify poor single loops and either rule them out as possible feedbacks or indicate the need for inner-loop equalization. Such loops may exhibit properties that make them candidates as accident causes if they are inadvertently opened or closed.

A systematic search for likely feedback loops has been accomplished using the 60 mph dynamics of Eqs. 1 – 4 and the driver describing function model. Compensatory closures (Fig. 3) were used. This is consistent with the "preview" concept when one considers that the effective controlled element dynamics, Y_c, include the geometry and kinematics of the external visual field as well as the dynamics of the vehicle. In essence, a guidance or control cue is perceived someplace in the visual field by the driver, and to the extent that the driver steers the vehicle to modify (e.g., follow or reduce) this cue in some way he is acting in a compensatory manner. The quality of the resultant closures was judged by

- The equalization required by the driver to provide a stable system consonant with the "adjustment rules"

- The closed-loop performance attainable as measured by estimated crossover frequency, stability margins, etc.

The results are summarized in Table I. The loop closures are illustrated by the system surveys of Figs. 4–6 for lateral deviation, path angle, and heading angle respectively.

The same closure criterion was used (where possible) for each of the loops in Table I in order to facilitate comparison. The same effective time delay was used where vehicle dynamics approximated K and K/s in the crossover region. A slightly larger time delay was used with K/s^2-like controlled elements.[6] In each case the time delay was larger than that given in Ref. 6 to account for some steering system lag.

TABLE I. SUMMARY OF SINGLE-LOOP SURVEYS

FEEDBACK	APPROXIMATE VEHICLE TRANSFER FUNCTION	DRIVER EQUALIZATION	ESTIMATED CROSSOVER FREQUENCY (rad/sec)	ESTIMATED DRIVER OPINION
Inertial Lateral Deviation $y_I \rightarrow \delta_w$	$\frac{405}{s^2}$	Large Lead	1.0	Poor; lead and low bandwidth
Heading Angle $\psi \rightarrow \delta_w$	$\frac{4.6}{s}$	Gain Only	1.5	Good
Path Angle $\gamma \rightarrow \delta_w$	$\frac{405}{s}$	Gain Only	1.0	Good; small lead helps
Lateral Velocity $v \rightarrow \delta_w$	-74	Lag-Lead	1.0	Fair; low bandwidth
Heading Rate $r \rightarrow \delta_w$	4.6	Lag-Lead	1.6	Good
Path Rate $\dot{\gamma} \rightarrow \delta_w$	405	Lag-Lead	1.5	Good
Lateral Acceleration at Driver's Head $a_y \rightarrow \delta_w$	400	Lag-Lead	1.0	Fair; low bandwidth and high sensitivity

Inertial lateral deviation is a necessary outer loop, but it requires lead equalization by the driver. This requirement for lead generation can be eliminated by the use of an inner loop. Inertial lateral deviation may be important as a single loop following perceptual transitions, e.g., when the driver suddenly loses his view down the road and is forced to steer on the basis of his lateral position in the lane. Such a regression should cause him to slow down so that the lower driver/vehicle system bandwidth

available when the driver must generate the lead would still enable him to follow any likely command input.

Both heading angle and path angle offer good closed-loop characteristics. Path angle control combines heading control (a good loop by itself) with body axis lateral velocity (sideslip), which is only fair taken alone. They can serve as outer loops if an intermittent trim loop is employed to occasionally reduce or reset the lateral deviations, or as inner loops to reduce the lead requirements for an inertial lateral deviation outer-loop system. The heading rate system is also good, as is path angle rate or curvature. The latter is the apparent curvature of the roadway ahead of the vehicle and has effective controlled element dynamics that are a pure gain in the region of crossover, implying driver lag equalization. The lateral acceleration at the driver's head is not a particularly good system because it is highly sensitive to driver gain variations.

Multiloop Structures

The good single-loop systems shown in Table I will all provide good performance in following a command input of that motion variable. None will do a very good job of following a path or trajectory which involves minimizing lateral position errors to stay in the center of the lane or roadway. Thus, the driver must augment his outer-loop structure with equalizing inner loops. Good multiloop system candidates are those which require little or no driver equalization (i.e., only gain plus time delay in each of the loops). Several possibilities have evolved to date, having the structures shown in Fig. 7. Three representative versions are summarized in Table II. The first entry relates to Fig. 7a; the other two have the structure of Fig. 7b.

The structure of Fig. 7a assumes that the driver operates on an estimated or projected lateral deviation error. Preview here is explicitly required, since this error is related to the lateral position the vehicle would have at a time T seconds (distance R) ahead of the vehicle if it continued along its current path (see Fig. 2). The perceptual preview, T, results in a pure lead equalization term in the effective controlled element dynamics. This, in turn, offsets the undesirable double integration form of the lateral deviation dynamics at low frequency. The Bode plot of Fig. 8 indicates that a T of 5 to 10 sec is sufficient. Values of T less than 5 sec (440 ft at 60 mph) are not as good because they do not compensate sufficiently for the

TABLE II. MULTILOOP SYSTEM CHARACTERISTICS

SYSTEM	INNER LOOP		OUTER LOOP		REMARKS
	Equalization	Crossover Frequency (rad/sec)	Equalization	Crossover Frequency (rad/sec)	
Inertial lateral deviation advanced in time	—	—	Gain only	1.2	Fine for straight roads. Difficult to define command input form and point of entry into the system. Sensitive to changes in preview.
Path angle (or path angle rate) plus inertial lateral deviation	Gain only	1.0	Gain only	0.53	Control not sensitive to changes in driver adaptation or driver attention. Outer loop can operate intermittently.
Heading angle (or heading rate) plus inertial lateral deviation	Gain only	1.5	Gain only	0.6	Control not sensitive to changes in driver adaptation or driver attention. Outer loop can operate intermittently.

inherent lags in the driver/vehicle system. This system structure provides a relatively high crossover frequency and good lateral position control on straight roads. Some conceptual and analytical difficulties arise when it is necessary to follow a curving roadway, e.g., a guidance law or scheme is needed to provide an appropriately advanced command input to compare with the projected lateral deviation.

The systems of Fig. 7b assume that the driver operates on some angular motion quantity plus lateral deviation as separate entities and combines them to produce a steer angle response. The inner loop in Fig. 7b could be any one of the four motion quantities indicated. This structure differs from the Fig. 7a model by having the present lateral deviation available to the driver for comparison with a desired command input and derivation of a position error. Bode plots for the lateral deviation loop closures with path angle and heading angle inner loops are given in Figs. 9 and 10, respectively. The effect on the outer-loop bandwidth of varying the inner loop gain is shown in Fig. 10.

Table II shows that both systems give adequate stability and reasonably good command-following. The stability of the systems are relatively insensitive to variations in the loop gains, and they are reasonably forgiving of momentary lapses in attention and do not require continuous control. Note that sideslip angles are usually relatively small, in which case heading angle and path angle are almost the same quantity. The larger value in Table II for the crossover frequency of the advanced-in-time system is an artifact of the loop closure criteria used. The closed-loop

dynamics of all three systems can be made very similar if the loop gain and preview times are appropriately adjusted. However, in making the estimates for path angle and heading angle, the inner loops are closed with relatively large stability margins, so as to be representative of conditions with possibly intermittent closures of the outer loop.

None of the systems shown in Fig. 7 include the two feedforward channels or the discrete response feedback loop of Fig. 1. The derived systems are, therefore, most appropriate for command-following or regulation tasks of a reasonably continuous nature with inputs that are more or less random appearing. The other types of response involving learned maneuvers, etc., are important and do dominate the driver's control activity in some phases of driving. However, closed-loop control is fundamental to many other phases of driving, and plays a key "take-over role" in circumstances where the learned maneuvers and patterned responses either cannot be structured or are suddenly destroyed for one reason or another.[12] The compensatory loops are also used in early phases of learning (the unskilled driver) and under conditions of extreme stress (the startled or confused driver). If the closed-loop systems won't or can't work, then there is little chance that the driver/vehicle system will function safely for long. Thus, understanding their form and operation gives the point of departure for either adding the other channels or switching to them as needed.

The systems in Table II do not exhaust the possible multiloop structures that can be concocted, but they are the only ones found to date which satisfy the guidance and control requirements for command-following and disturbance regulation with good performance, insensitivity to variations in the driver's dynamic adaptation, good predicted subjective opinion from the driver, etc. Further, they are not inconsistent with available perceptual data obtained from driving experiments on the highway. The principal distinctions between them lie in the way in which they degrade when the nominally good structure is disturbed.

Correlation With Experimental Observations

The identification of preferred driver/vehicle loop closures and sensory cues from a control theory approach provides a new framework for reviewing past perceptual experiments.

Gordon[13] attempted to define the driver's visual input, using a helmet-mounted aperture for one eye comprised of a tube 3.5 in. long with a variable

diameter of 1 in. or less. The other eye was masked, and a camera was mounted on the helmet. The task was to drive along a narrow, winding, road. Most records showed continuous visual shifts forward to the limit of the visible road and then backward toward the vehicle, suggesting that the driver may be looking down the road to obtain angle information, and then near the vehicle to sample lateral position. About 60 percent of the time the driver was looking farther than 150 ft ahead of the vehicle, while about 80 percent of the time he was looking more than 100 ft ahead. At the test speeds of about 15 mph these are preview times of 7.5 and 5 sec, respectively, which correlate well with the desirable times in Fig. 8. Small aperture viewing caused stress; peripheral cues may be important.

More recently, Gordon[14] expands on the "streamer theory" (after Calvert[15]), which states in essence that the driver perceives motion from objects in the visual field streaming across his field of view and emanating from a central focus. Gibson[16], on the other hand, asserts that this "focus of expansion" provides the directional cue rather than the flow of the velocity field emanating from the focus. Regardless of how the driver senses the direction to this focus, this cue is precisely path angle. Schmidt[17] notes in agreement, "the driver recognizes the movement of the car from the apparent flow or streaming of the objects in the visual field..."

Biggs[18] concludes that central vision is occupied with the detection of obstacles in the immediate path while peripheral vision is employed in tracking the "guideline" or dividing line. He notes that "the guideline is seen near the vehicle, and its lateral motion provides the dominant directional cue."

Crossman[19] hypothesized a family of more or less distinct control systems or modes of driver control, and tested them by comparing driving performance in a fixed-base simulator with that obtained under actual driving conditions on the highway. He showed better performance on the highway, apparently due to simulator shortcomings. His results suggest that the driver responds to heading (or path angle) errors when they exceed a threshold level.

Although recent, none of these references specify the alternative closed-loop structures appropriate to various driving situations. They do support, however, the theory of driver control derived herein. While there are apparent differences between these experimental results, some semblance

of order is obtained by viewing them within this guidance and control structure, and the operational entities extracted from the perceptual cues do provide the prescribed feedbacks.

CONCLUSIONS

Guidance and control principles coupled with empirically based models for driver dynamics can be used to evolve alternative descriptions of driver steering control. The resultant system models provide an engineering tool useful in efforts to improve vehicle dynamics, enhance driver perception and control, and optimize the roadway environment from the standpoint of safety and service volume.

REFERENCES

1. Weir, D. H., et al, Dynamics of the Automobile Related to Driver Control, Systems Tech., Inc., TR 157-1, July 1966. Also, SAE Paper No. 680194.
2. Weir, D. H., et al, A Theory for Driver Steering Control of Motor Vehicles, 47th Annual Mtg. of Highway Res. Board, Wash., D. C., Jan. 1968.
3. Segel, L., On the Lateral Stability and Control of the Automobile as Influenced by the Dynamics of the Steering System, ASME Paper 65-WA/MD-2.
4. McRuer, D. T., et al, "The Human Operator as a Servo System Element," J. Franklin Inst., Vol. 267, Nos. 5 and 6, 1959.
5. Wasicko, R. J., et al, Human Pilot Dynamic Response in Single-Loop Systems with Compensatory and Pursuit Displays, AFFDL-TR-66-137, Dec. 1966.
6. McRuer, D., et al, Human Pilot Dynamics in Compensatory Systems, AFFDL-65-15, July 1965.
7. McRuer, D., et al, "Manual Control of Single-Loop Systems," J. Franklin Inst., Vol. 238, Nos. 1 and 2, 1967.
8. Magdaleno, R. E., et al, Small Perturbation Dynamics of the Neuromuscular System in Tracking Tasks, Systems Tech., Inc., TR 154-1, Oct. 1967.
9. Stapleford, R. L., et al, Pilot Describing Function Measurements in a Multiloop Task, NASA CR-542, Aug. 1966.
10. McDonnell, J. D., Pilot Rating Techniques for the Estimation and Evaluation of Handling Qualities, AFFDL-TR-68-76.
11. McRuer, D. T., et al, New Approach to Human-Pilot/Vehicle Dynamic Analysis, AFFDL-TR-67-150, Oct. 1967.
12. Weir, D. H., et al, Conceptualization of Overtaking and Passing on Two-Lane Rural Roads, Vol. III: Driver Control, Systems Tech., Inc., TR 1-193, Dec. 1967.
13. Gordon, D. A., "Experimental Isolation of Drivers' Visual Input," Public Roads, Vol. 33, No. 12, Feb. 1966.
14. Gordon, D. A., "Perceptual Basis of Vehicular Guidance," Public Roads, Vol. 34, No. 3, Aug. 1966.
15. Calvert, E. S., "Visual Judgments in Motion," J. Inst. of Navigation, Vol. 7, No. 3, 1957.

16. Gibson, J. J., The Perception of the Visual World, Houghton Mifflin, Boston, 1950.

17. Schmidt, I., et al, Visual Considerations of Man, the Vehicle, and the Highway, SAE SP-279, Mar. 1966.

18. Biggs, L., "Directional Guidance of Motor Vehicles—A Preliminary Survey and Analysis," Ergonomics, Vol. 9, No. 3, May 1966.

19. Crossman, E. R. F. W., et al, Steering Performance of Automobile Drivers in Real and Contact-Analog Simulated Tasks, Human Factors Soc. 10th Annual Mtg., Oct. 1966.

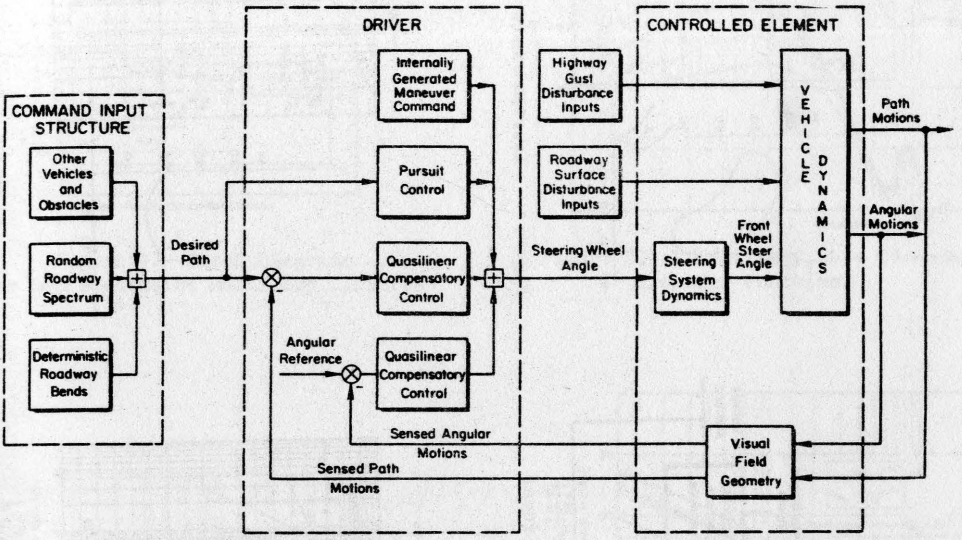


Figure 1. Topology of the Closed-Loop Control Structure

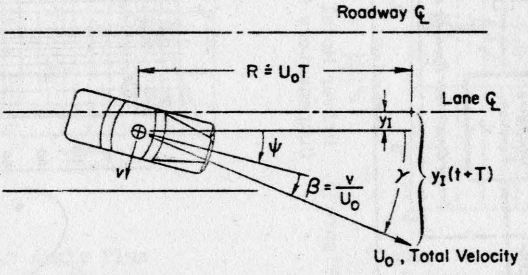


Figure 2. Motion Quantities for Directional Control

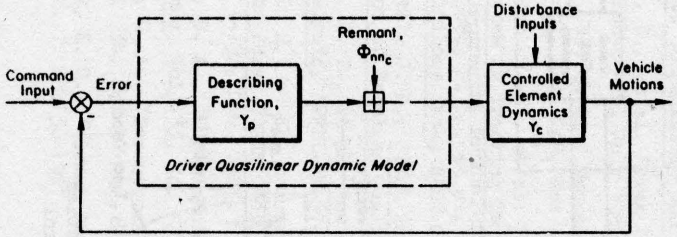


Figure 3. Driver Model for Quasi-Linear Compensatory Control

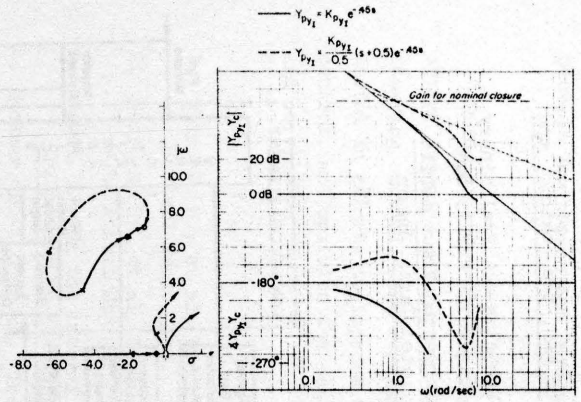


Figure 4. Inertial Lateral Deviation Loop Closures

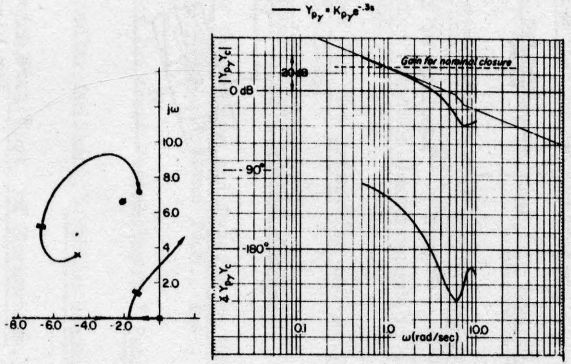


Figure 5. Path Angle Loop Closure

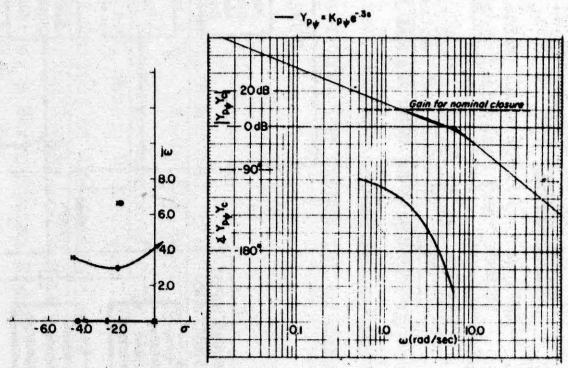


Figure 6. Heading Angle Loop Closure

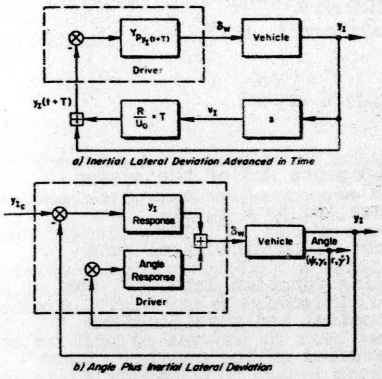


Figure 7. Block Diagrams for Representative Multiloop Closures

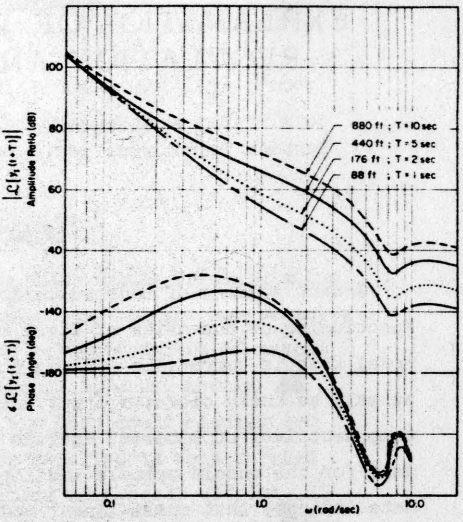


Figure 8. Effective $y_T(t + T) \rightarrow \delta_w$ Transfer Function

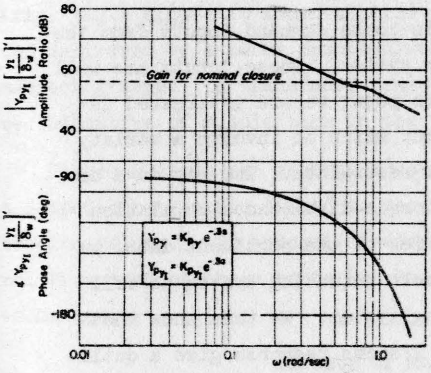


Figure 9. Path Angle Plus Lateral Deviation Closure

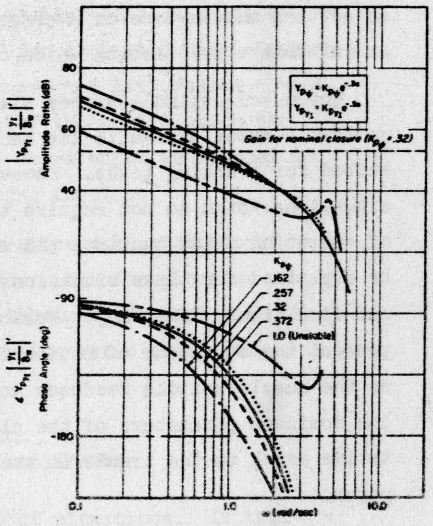


Figure 10. Heading Angle Plus Lateral Deviation Closure

A CLOSED-LOOP NEUROMUSCULAR SYSTEM EXPLANATION OF FORCE-DISTURBANCE REGULATION AND TREMOR DATA*

R. E. Magdaleno and D. T. McRuer
Systems Technology, Inc., Hawthorne, California, U. S. A.

INTRODUCTION

Recent studies of neuromuscular system components and of the system functioning of ensembles of these components has led to a quasi-linear model that is compatible with an enormous amount of physiological data as well as human operator input/output describing function data.[1,2] The component data of interest include recent anatomical and physiological data for the muscle spindle and input/output studies of the muscle. These data indicate that quasi-linear models can describe the basic behavior of these two elements for small perturbations about an operating point.[2] This model contains three key developments: (1) the variation in muscle system parameters as a function of average muscle tension; (2) the role of the muscle spindle, both as an adaptive equalization element and in its effect on setting muscle tone or average tension; and (3) the spindle feedback of an internal muscle length which constrains some dominant closed-loop roots.

The muscle spindle system also serves to route command inputs from the central nervous system to the neuromuscular system. These inputs are well suited for tracking tasks. However, in this paper we are interested in situations which do not require tracking, but which do involve a variety of operating point tensions and manipulator restraints. The key data to be explained for these situations are for torque-disturbance regulation and limb tremor frequency. Following a review of the pertinent data, we present the equations of motion for the muscle actuating system as well as the muscle spindle feedback and actuating system. We then show that the dominant parameters of the closed-loop systems response give a qualitative match to the trends in the data reviewed at the beginning of the paper.

---

*This paper includes research efforts supported by the Ames Research Center, Man-Machine Integration Branch, NASA, Moffett Field, California, under Contract NAS2-2824.

HUMAN OPERATOR TORQUE-DISTURBANCE AND TREMOR DATA

In this section we shall review torque-disturbance and limb tremor data obtained for a variety of operating point tensions and manipulator restraints.

Torque-Disturbance Regulation

The effect of average voluntary muscle tension on the wrist-twisting (supination/pronation) response to torque-disturbance inputs has recently been investigated.[3-7] In one experiment[4] the relative amount of muscle tension was inferred from a sphygmomanometer cuff attached around the forearm. This was displayed to the subject who then could readily set the reading to any one of five levels. These experiments were carried out using irregularly spaced mechanical impulses delivered, without warning, by a pendulum. The manipulator restraint consisted of an inertia several times larger than that of the arm, and the subject was asked to resist the perturbing influence of the pendulum-produced disturbance on the load.

The transient responses obtained in these tests resembled those of a unit numerator second-order system weighting function with light damping (Fig. 1a). Figure 1b shows similar results[6] where the disturbance torque was delivered by a motor attached to the shaft of the handle which the subject grasped. A transducer sensed the grip force which is assumed to be indicative of muscle tone in the participating muscles.

Figure 2 shows the general location of the upper pole position of a damped second-order fitted to the transient response for a number of tension values.[4] In general, increasing mean tension increases the natural frequency of these roots but leaves the damping ratio relatively unchanged.

Effects of Muscle Tension and Manipulator Restraints on Limb Tremor

Limb tremor has been observed in a variety of situations. It typically appears in the time history as an oscillation or in the power

spectral density as a peak. Tremor frequencies are often in the range of 6-10 Hz, as in the analysis of finger velocity[8] and with different load masses.[9] Generally, muscle tremor frequencies depend on both the manipulator restraint and muscle tension.[10-12] In Fig. 3 forearm flexor tremor frequency is plotted as a function of the mean tension at the wrist.[10] This tension was exerted against a spring (attached to the wrist) which was oriented parallel to the upper arm. For each spring, increasing the tension causes a slight increase in tremor frequency, with a leveling off at the higher tensions. However, an increase in spring rate produces an incremental increase in tremor frequency which is essentially independent of tension.

The effect of inertia on limb tremor was found in a study where the task was to track a ramp input to a pursuit display.[11] The subject gripped a handle in his fist and used wrist rotation about the forearm axis to generate his response. Four different load inertias were used. Muscle tension was inferred from handle grip tension as sensed from a rubber pneumatic balloon and pressure-voltage transducer. The dependence of tremor frequency on inertia and inferred muscle tone is given in Fig. 4. For each inertia there is an increase in frequency as tone increases, whereas for constant tone an increase in inertia causes a decrease in frequency.

Additional results[11] indicate that the tremor frequency of one hand is independent of the other. Specifically, if one hand is controlling a large inertia without pressure, then its tremor frequency can be much less than that of the other hand if it is controlling a small inertia with great pressure. This suggests that the oscillation conditions are strongly dependent on the load and further that each neuromuscular system is independent of the other.

CLOSED-LOOP NEUROMUSCULAR SYSTEM DYNAMICS

In this section we describe the joint action of the component models (muscle actuation and spindle feedback and actuation subsystems) as they contribute to the neuromuscular system model. This will be used to describe the trends in the overall input/output human operator data discussed in the last section.

Neuromuscular System Block Diagram for Small Perturbations

A combination block and schematic diagram of the muscle/spindle system for small perturbations about an operating point is shown in Fig. 5. Because the schematic represents perturbation operations about steady-state operating points, all the signals indicated vary about zero mean values, and thus the agonist/antagonist relationships are subsumed in a one-sided composite. For this paper the details of the muscle spindle model and muscle models are not required — only their input/output relations will be needed. (For the details see Ref. 2.)

The internal inputs to the system are: (1) command, γ_c, and bias, γ_b, which are aspects of the static and dynamic gamma muscle spindle fibers; (2) the alpha motor neuron commands, α_c. The external input is a force, F, applied to the manipulator. A key feature is the spindle feedback of x_1 (an internal muscle length) rather than x the limb length. Not shown in Fig. 5 is any consideration of remnant sources necessary to account for system behavior not described by the quasi-linear elements.

The spindle output perturbation firing rate is summed with an alpha motor neuron command, with the result, after conduction and synaptic delays, being an incremental alpha motor neuron firing rate, Δf_α, about the operating point, f_{α_o}. This in turn perturbs the muscles and manipulator, ultimately giving rise to limb rotation, which is sensed by the spindle ensemble. The steady-state output of the muscle spindles (not shown in Fig. 5) provides the steady-state muscle firing rate, f_{α_o}.

The muscle model characteristics consist of the series elastic component, K_e, connected in series with the contractile component characteristics given by the force source, $C_f \Delta f_\alpha$, spring, K_m, and damper, B_m. (K_e also contains tendon compliance.) Since the limb and manipulator inertias are in parallel, the two are lumped together in the single effective inertia, M. A load spring, K_F, is also present.

The muscle effective damping, B_m, spring rate, K_m, and sensitivity, C_f, are set by the operating point muscle tension, P_o, which is due to the steady-state output of the muscle spindle. B_m, K_m, and C_f increase as P_o increases (approximately linear functions).[2]

We shall first find the open-loop equations of motion connecting the muscle system inputs (Δf_α and F) and its responses (x and x_1). Then we shall apply the spindle feedback dynamics and solve for the closed-loop response (x) to the system inputs (F, γ_c, and α_c).

The Laplace transformed equations of motion for the open-loop limb/manipulator dynamics are given in matrix form by

$$\begin{bmatrix} (Ms^2 + K_F + K_e) & -K_e \\ -K_e & (B_m s + K_m + K_e) \end{bmatrix} \begin{bmatrix} x \\ x_1 \end{bmatrix} = \begin{bmatrix} F \\ C_f \Delta f_\alpha \end{bmatrix} \tag{1}$$

The characteristic equation is

$$\Delta = MB_m \left[s^3 + \left(\frac{K_m + K_e}{B_m} \right) s^2 + \left(\frac{K_F + K_e}{M} \right) s + \frac{K_F(K_m + K_e) + K_m K_e}{MB_m} \right] \tag{2}$$

A number of special cases have been investigated, ranging from an isometric manipulator ($K_F \rightarrow \infty$) to a large-inertia, no-spring case.[2] It was found that the approximate factors of the third-order characteristic equation indicate that the real root is the same as that for an isometric manipulator. Using this, the approximate factors to Eq. 2 become

$$\Delta \doteq MB_m \left(s + \underbrace{\frac{K_m + K_e}{B_m}}_{\text{(A)}} \right) \left[s^2 + \underbrace{\frac{B_m}{M} \left(\frac{K_e}{K_m + K_e} \right)^2}_{\text{(B)}} s + \frac{1}{M} \left(K_F + \frac{K_m K_e}{K_m + K_e} \right) \right] \tag{3}$$

which applies as long as (A) is much greater than (B). This approximation holds for large inertia and large K_e. Generally, K_e, due to tendon elasticity, is large relative to K_m.

The muscle spindle typically has a lead/lag response to length changes. The perturbation firing rate, Δf_α, depends on the spindle feedback of x_1 and on α_c and γ_c as[2]

$$\Delta f_\alpha = [\alpha_c + G_s(\gamma_c - x_1)] e^{-\tau_\alpha s} \tag{4}$$

where $G_s = K_s \dfrac{(s + 1/T_K)}{(s + 1/aT_K)}$

and τ_α is the net effect of synaptic and conduction time delays, γ_c is the effective command due to gamma fiber stimulation, $1/T_K$ is the lead break frequency, and $1/aT_K$ is the lag break frequency. Now inserting Eq. 4 into Eq. 1 and collecting terms that multiply x_1 yields the complete system equations for the muscle/manipulator/muscle spindle system.

$$\begin{bmatrix} (Ms^2 + K_F + K_e) & -K_e \\ -K_e & (B_m s + K_m + K_e) + C_f G_s e^{-\tau_\alpha s} \end{bmatrix} \begin{bmatrix} x \\ x_1 \end{bmatrix} = \begin{bmatrix} F \\ (\alpha_c + G_s \gamma_c) e^{-\tau_\alpha s} \end{bmatrix} \tag{5}$$

The limb output response to the various forcing function terms, F, α_c, and γ, is

$$x = \frac{\left[(B_m s + K_m + K_e) + C_f G_s e^{-\tau_\alpha s}\right] F + \left[K_e(\alpha_c + G_s \gamma_c) e^{-\tau_\alpha s}\right]}{\Delta + C_f G_s (Ms^2 + K_F + K_e) e^{-\tau_\alpha s}} \tag{6}$$

When spindle feedback is present, the closed-loop poles are modified from the zeros of Eq. 3. However, the spindle feedback, G_s, also modifies the numerator dynamics of the limb response to a force-disturbance input. This will be significant in the interpretation of the torque-disturbance input data. The effect of the spindle feedback on the limb response to command inputs is primarily through the denominator roots.

The two situations of interest in this paper are:

1. An impulse force-disturbance input delivered when the subject's agonist/antagonist muscles are in tension against each other, and the manipulator is loaded with a large inertia (no spring restraint). These impulses are delivered at random times and directions, thus the subject's voluntary actions are largely confined to setting the muscle tension so that the resulting response reflects the properties of the x/F transfer function, i.e., α_c and γ_c are zero. The subject's task is merely to set muscle tone and let the transient run its course.

2. Limb tremor under various tension conditions where, again, α_c and γ are zero except for ambient noise-like perturbations (also for this case, $F = 0$).

Force-Disturbance Input (No Spring, Large Inertia)

The effect of the spindle feedback on the x/F response can be found by noting that the zeros are given by the roots of the numerator of Eq. 6 found from the Numerator Closure as

$$1 + \frac{C_f G_s e^{-\tau_\alpha s}}{B_m s + K_m + K_e} = 0 \qquad (7)$$

and the poles are given by the roots of the denominator of Eq. 6 found from the Denominator Closure as

$$1 + \frac{C_f G_s (Ms + K_F + K_e) e^{-\tau_\alpha s}}{\Delta} = 0 \qquad (8)$$

where for this manipulator the roots of Δ are approximated by Eq. 3 with $K_F = 0$ which yields

$$\Delta \doteq MB_m \left(s + \frac{K_m + K_e}{B_m}\right)\left[s^2 + \frac{B_m}{M}\left(\frac{K_e}{K_e + K_m}\right)^2 s + \frac{K_m}{M}\left(\frac{K_e}{K_e + K_m}\right)\right] \qquad (9)$$

Both the numerator and denominator of the closed-loop function x/F are functions of the gain parameter, $C_f \kappa_s$. The joint movement of the numerator and denominator singularities as a function of this loop gain ($C_f \kappa_s$) can be demonstrated by treating both the numerator and the denominator as loop closures. Then both closed-loop poles and closed-loop zeros will move as loop gain is changed. Root locus sketches of the Denominator and Numerator Closures are given in Fig. 6a and Fig. 6b, respectively. These two closures have common real-axis poles and time delay terms (the latter are not shown explicitly since they are at higher frequencies). Only the Denominator Closure will have quadratic roots (at high frequency). The closed-loop low frequency real root will be approximately the same in

both closures. Thus the closed-loop system describing function, x/F, is dominated by a second-order mode which has an undamped natural frequency which is a function of the loop gain (i.e., operating-point tension dependent) as well as the external mechanical element (inertia). This is compatible with the Ref. 4 results of Fig. 2. However, the degree of pole/zero cancellation will be affected by the proximity of the quadratic terms. Some of the different responses are shown in Fig. 1. In Fig. 1b the response seems to contain some of the first-order pole, whereas in Fig. 1a the response is nearly a pure second-order transient.

Now examine the consequences of assuming that the spindles feed back x instead of x_1 (in Fig. 5 let $K_e \rightarrow \infty$). Carrying this through the equations reveals that the numerator of the x/F response no longer contains $C_f G_s$. Now the pole/zero near-cancellation will not take place, and the transient response will appear third-order. This will be more evident at the higher tensions since the increased loop gain drives the $1/T_K'$ root to lower frequency away from the numerator singularity at $(K_e + K_m)/B_m$. This trend is just the opposite of that in Fig. 1b, which becomes more like a second-order system at higher tensions.

Tremor

A number of theories have been advanced to account for tremor oscillations. These include: (1) periodically fluctuating activity in supraspinal centers; (2) rhythmicity arising from neuronal circuits within the spinal cord itself; (3) resonance of the moving parts of the limbs; (4) oscillations arising from the stretch-reflex loop itself.

At present, the evidence very much favors the last hypothesis, since (1) in tabetic patients with loss of dorsal root fibers the tremor frequency (6-10 Hz) peak is absent;[16] (2) there are no substantive reports indicating any correlation of EEG activity (alpha rhythm) with limb tremor; (3) tremor disappears with section of the dorsal roots;[16, 17] (4) there is no evidence of inherent rhythmicity of spinal alpha or gamma motor neurons in the absence of sensory input. Also, tremor occurs primarily under

conditions where fine control of position is involved; much less tremor is produced under rest conditions or when a gross, rapid ("ballistic") type of movement is executed.[18] This also tends to support the stretch-reflex (closed-loop) explanation.

The "tremor peak" in the frequency spectrum is said by some[13] to be indifferent to changes in the load, but the data[10-12] of Figs. 3 and 4 clearly contradict this, and in addition indicate that the peak frequency is dependent on the mean tension generated, spring rate, and inertia.

The data cited above and earlier indicate that any neuromuscular system model must be compatible with ambient noise excitation of an almost neutrally stable mode near 6-11 Hz which varies with inertia, spring rate, and tension. The closed-loop system for these situations in which tremor was measured is described by Eq. 6 with $F = 0$. In addition, α_c and γ_c are nominally zero, although they will exhibit random fluctuations since the human operator cannot generate a perfectly constant tension. These fluctuations will excite the lightly damped closed-loop pole (Fig. 6a) for the case of inertia restraint and no spring. Increases in inertia will decrease the frequency of this mode, whereas increases in tension, by increasing loop gain, will drive the roots closer to the zeros on the $j\omega$-axis. Both of these trends are compatible with the data in Fig. 4.

For the data given in Fig. 3 a load spring, K_F, was present. Thus, using Eq. 3 for the approximate roots of Δ and inserting in Eq. 6, yields the root locus in Fig. 7. The load spring has increased the frequency of the quadratic poles and zeros. Tension increases (by increasing loop gain) drive the mode into the zeros, but now the instability frequency is less than that of zero due to the phase shift from the time delay $e^{-\tau_\alpha s}$. The spread between the zero and the instability frequency is likely to be greater for the stiffer spring rates since this raises the zero and the pole to higher frequency where the loop time delay (Eq. 6) has more effect. At higher frequencies the time delay causes the root locus departure from the pole to occur at a lower angle, i.e., it tends to proceed more directly to the $j\omega$-axis. Thus increases in spring rate yield an instability frequency that is less than the frequency of the

zero, and this difference increases at the higher frequencies. This effect can be seen in the flattening of the curves in Fig. 3 for the larger spring rates which indicates that the root is near a limit. Attempting to put the above explanation on a quantitative basis by estimating the inertia involved in the Fig. 3 data, using reasonable values of K_e, and then calculating the effect of K_F, yields predicted ω values for the quadratic zeros of between 7 and 14 Hz. These are "infinite loop gain" tremor frequencies, as compared to the actual tremor frequency range from 7 to 11 hz given in Fig. 3.

A time delay on the order of 25 ms is compatible with nerve conduction time from periphery to spinal cord to periphery.[14,15] Such a value would contribute 90 deg phase shift at 10 Hz, leaving 90 deg to be contributed by the other loop dynamics to produce the 180 deg instability frequency. The loop gain function (Eq. 8) can be compatible with phase shifts of 90 deg near 10 Hz.* Thus the tremor data is compatible with ambient noise excitation of a closed-loop neuromuscular system mode.

Finally, the power spectra of finger tremor[8] exhibited a second peak in the range 17-25 Hz in addition to the usual peak in the range 6-10 Hz. This data indicates a higher frequency mode for which a likely source is another lightly damped mode pair due to a real root proceeding in from the time delay singularity at minus infinity, combining with the root at $1/aT_K$ (Fig. 7) and then nearing neutral stability at frequencies above the quadratic pole/zero combination.

CONCLUSIONS

The data and analysis presented in this paper indicate that simple quasi-linear models can describe the small perturbation behavior of the major neuromuscular system components (muscles and muscle spindles) for torque disturbance regulation and limb tremor.

---

*For the special case of an isometric manipulator ($K_F \rightarrow \infty$), the quadratic poles and zeros tend to infinity, leaving only the two real poles and a zero. If the frequency of these singularities is smaller than 10 Hz, then they would contribute approximately 90 deg phase lag at 10 Hz.

The more specific conclusions are:

- Changes in the transient response to a torque impulse input as a function of muscle tension (large-inertia, no-spring manipulator) support the model's muscle spindle feedback of an internal muscle length. The effect of this feedback is to produce a closed-loop numerator zero which approximately cancels a denominator root since both are influenced by operating point tension. The resulting dominant second system response closely describes the torque disturbance data.
- Variations in muscle tremor frequency for various spring, inertia, and average muscle tension values support a closed-loop neuromuscular system model. The trends and limits of the tremor data provide further evidence of spindle feedback of an internal muscle length rather than a length directly proportional to limb length. The series elastic element coupled with limb/manipulator inertias produces zeros which constrain the closed-loop characteristic equation roots and provide an upper limit for tremor frequency.

SYMBOLS

$B_m(P_o)$	Equivalent damper for contractile component of muscle system
C_f	Constant of proportionality for muscle model force source
f_{α_o}	Operating point average firing rate for muscle
Δf_α	Incremental alpha motor neuron firing rate
F	External force input
G_s	Spindle model describing function
$j\omega$	Imaginary part of the complex variable, $s = \sigma + j\omega$
K_e	Series elastic component of muscle
K_F	Load spring
K_m	Equivalent spring gradient for contractile component of muscle system
L_o	Operating point length of muscle
ms	Millisecond
M	Limb plus load inertia
M	Meters

N	Newtons
P_o	Operating point tension in agonist or antagonist muscle
s	Complex variable, $s = \sigma + j\omega$; Laplace transform variable
$\frac{1}{T_K}$	Open-loop spindle model zero
$\frac{1}{aT_K}$	Open-loop spindle model pole
x	Limb length in muscle/manipulator model
x_1	Internal muscle length
α_c	Alpha motor neuron command
γ	Gamma command following gamma motor neuron delay
γ_b	Gamma bias input to spindle model
γ_c	Gamma command input to spindle model
γ_d	Gamma input due to dynamic gamma motor neuron
γ_s	Gamma input due to static gamma motor neuron
Δ	Characteristic equation
ζ_N	Damping ratio of second-order component of the neuromuscular system
K_s	Gain of spindle describing function, G_s
σ	Real part of the complex variable $s = \sigma + j\omega$
τ_α	Net time delay in the alpha motor neuron pathway
ω_N	Undamped natural frequency of second-order part of the neuromuscular system
$\doteq$	Approximately equal to
$(\ )^\circ$	Degrees
∞	Infinity

REFERENCES

1. McRuer, D. T., R. E. Magdaleno, and G. P. Moore, "A Neuromuscular Actuation System Model," Third Annual NASA-University Conference on Manual Control, NASA SP-144, 1-3 Mar. 1967, pp. 281-304.

2. Magdaleno, R. E., G. P. Moore, and D. T. McRuer, Small Perturbation Dynamics of the Neuromuscular System in Tracking Tasks, forthcoming NASA Contractor Report.

3. Young, Laurence R., and Lawrence Stark, Biological Control Systems — A Critical Review and Evaluation, Developments in Manual Control, NASA CR-190, Mar. 1965.

4. Okabe, Y., H. E. Rhodes, L. Stark, and P. A. Willis, Transient Responses of Human Motor Coordination System, MIT Res. Lab. Elec. Quar. Prog. Rept. No. 66, July 1962, pp. 389-395.

5. Houk, James Charles, Jr., A Mathematical Model of the Stretch Reflex in Human Muscle Systems, Master of Science Thesis, MIT, 1963.

6. Sun, H. H., B. A. Eisentein, and H. Bomze, "Dynamic Model for Hand Motor Coordination System," Eng. in Med. and Biology, Proc. 19th Annual Conference, 1966.

7. Stark, Lawrence, "Neurological Feedback Control Systems," Advances in Bioengineering and Instrumentation, Chapter 4, Fred Alt (ed.), Plenum Press, New York, 1966, pp. 289-385.

8. Wyatt, R. H., Jr., "A Study of Power Spectra Analysis of Normal Finger Tremors," IEEE Trans., Vol. BME-15, No. 1, Jan. 1968, pp. 33-45.

9. Holliday, A. M., and J. W. T. Redfearn, "An Analysis of the Frequencies of Finger Tremor in Healthy Subjects," J. Physiol., Vol. 134, 1956, pp. 600-611.

10. Robson, J. G., "The Effect of Loading on the Frequency of Muscle Tremor," J. Physiol., 1959, pp. 29P.

11. Gydikov, A., "Sampling with Adjustable Frequency in the Hand Movement Control System," IEEE Trans., Vol. HFE-8, No. 2, June 1967, pp. 135-140.

12. Sutton, G. G., and K. Sykes, "The Variation of Hand Tremor with Force in Healthy Subjects," J. Physiol., Vol. 191, 1967, pp. 699-711.

13. Lippold, O. C. J., J. W. T. Redfearn, and J. Vuco, "The Rhythmical Activity of Groups of Motor Units in the Voluntary Contraction of Muscle," J. Physiol., Vol. 137, 1957, pp. 473-487.

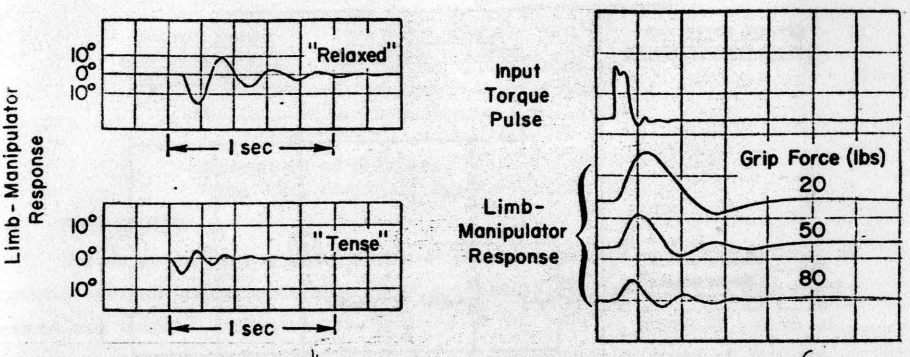
14. Margaria, R., T. Gualtierotti, and D. Spinelli, "Effect of Stress on Lower Neuron Activity," Exp. Med. Surg., Vol. 16, No. 2-3, June-Sept. 1958.

15. Dewhurst, D. J., "Neuromuscular Control System," IEEE Trans., Vol. BME-14, No. 3, July 1967, pp. 167-170.

16. Holliday, A. M., and J. W. T. Redfearn, "Finger Tremor in Tabetic Patients and Its Bearing on the Mechanism Producing the Rhythm of Physiological Tremor," J. Neurol. Neurosurg. Psychiat., Vol. 21, 1958, pp. 101-108.

17. Lippold, O. C. J., J. W. T. Redfearn, and J. Vuco, "The Influence of Afferent and Descending Pathways on the Rhythmical and Arrhythmical Components of Muscular Activity in Man and the Anaesthetized Cat," J. Physiol., Vol. 146, 1959, pp. 1-9.

18. Gibbs, C. B., "Function of Limb Tremor," Nature, Vol. 190, 6 May 1961.



(a) Pendulum Input[4] (b) Torque Motor Input[6]

Figure 1. Responses to a Torque Disturbance for Various Tension Levels

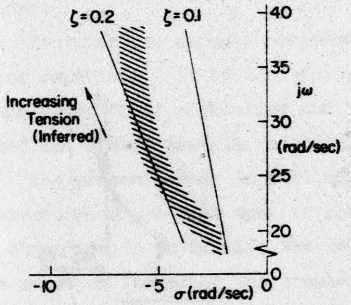


Figure 2. Upper Pole Location of a Damped Second Order Fitted to Transient Responses[4]

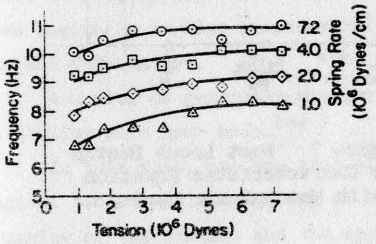


Figure 3. Effects of Muscle Tension and Spring Rate on Forearm Flexor Tremor[10]

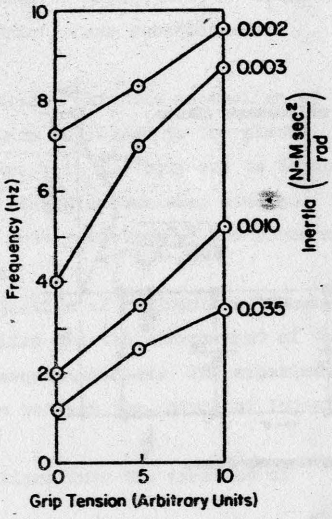


Figure 4. Effects of Inertia and Grip Tension on Limb Tremor During Constant Velocity Wrist Rotation[11]

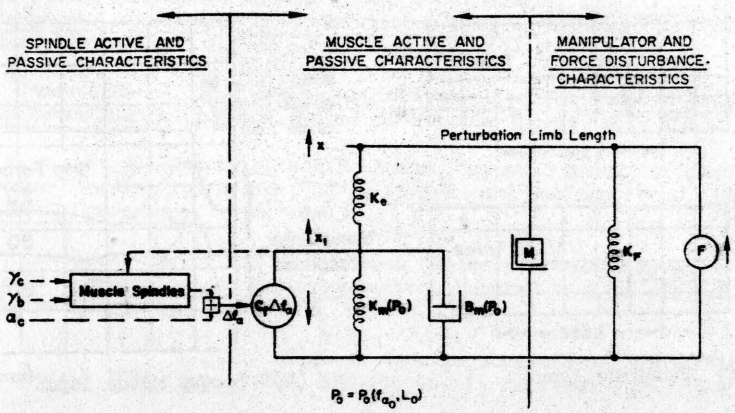


Figure 5. Combination Block and Schematic Diagram for the Muscle/Manipulator/Muscle-Spindle System[2]

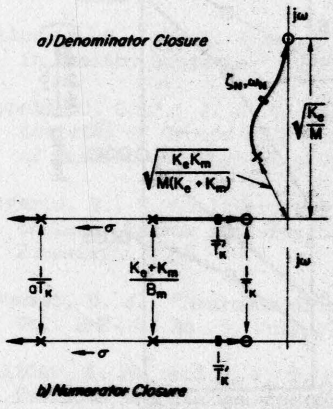


Figure 6. Root Locus Sketch for Denominator and Numerator Roots for Force-Disturbance Input

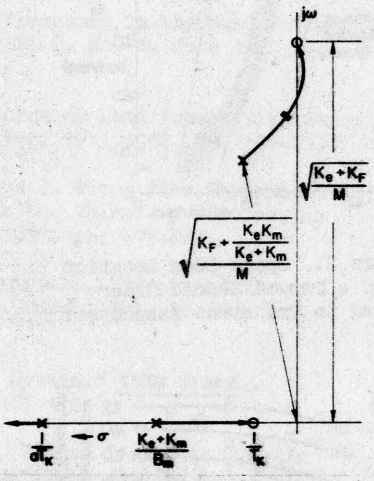


Figure 7. Root Locus Sketch for Characteristic Equation with Manipulator Spring

A STUDY OF THE DYNAMICS OF PLASMA PROTEIN METABOLISM

C.D. Barr, E.R. Carson, L. Finkelstein
Department of Automation Engineering,
The City University, London E.C.1., U.K.
E.A. Jones,
Department of Medicine,
Royal Free Hospital, London W.C.1., U.K.

1. INTRODUCTION

Recent years have seen an increase in the application of the methods of dynamic systems analysis to biological problems. The reasons for this development are twofold: firstly, the methods of the systems engineer can be of assistance in biological investigations; and secondly, a study of biology may well yield important clues to the principles of organisation of complex systems in general. These considerations stimulated the present study in which the methods of engineering dynamic analysis were applied to a metabolic process.

Proteins are organic compounds which are essential constituents of all living organisms. It is therefore most important to understand the dynamic mechanism of their metabolism and the way in which these substances are formed and broken down in the body.

The present study is part of medical research into the metabolism of liver-produced proteins such as albumin which are delivered to the plasma. The objective is to clarify the metabolic pathways in the body and to determine rates of formation and breakdown. Urea metabolism was also studied. The processes are essentially dynamic and the interpretation of the observed data requires an understanding of the dynamics.

The experimental method involved the injection of radioactive tracers into the body at various points of the metabolism and the measurement of the label at other points at intervals over a subsequent period. The experiments gave basically impulse responses of the system between the points of injection and observation.

A number of generally similar investigations into the dynamics of metabolism have been made.[1-4]

In this work, use was made of a mathematical model formulated from biological considerations, giving in our view a clearer comprehension of the dynamics of the process and the approximations involved. Also the methods of fitting responses to the model will make fuller use of the experimental data, since the whole - and not only, say, the latter- part of the response curves was considered.

2. FUNCTIONAL MODEL OF THE METABOLIC SYSTEM

As a first step to the formulation of the mathematical model of the

metabolic system it was necessary to arrive at a descriptive functional model.

The basic concept of this model is that of the compartment or pool, that is the quantity of a substance present in the body or a particular part of the body. Examples of pools are the total amount of albumin inside the blood vessels and the total amount outside the blood vessels.

From biological considerations, the following is considered to represent the system of albumin and urea metabolism, starting from labelled carbonate injected into the blood. (Fig 1)

After injection of the labelled carbonate into the bloodstream, most is lost through the lungs and some through the gut, and the remainder of the labelled carbon passes into the liver where both albumin and urea are synthesised.

In the liver many complex biochemical reactions take place in the course of which labelled carbon is incorporated into many intermediary metabolic cycles including several amino acids. Labelled carbon entering these pathways is either lost from the system or after a variable time interval is returned to the liver.

The amino-acid of specific interest in this study was arginine. When the labelled arginine[5] is produced in the liver it is used to synthesise the following four groups of labelled substances: (1) liver cell proteins (2) albumin (3) urea (4) other liver produced plasma proteins.

The proteins incorporated into the structure of the liver cell have a varying life span. Their catabolism liberates arginine into the liver arginine pool.

The labelled urea and albumin then pass into the plasma. A time delay exists between the appearance of urea and that of albumin in the plasma. One explanation is that albumin following synthesis is passed to some sort of storage pool located in the liver, before entering the plasma.

Once in the plasma the albumin can pass freely to and from an extravascular pool. Also the albumin may be catabolised at sites which are probably in or near the plasma compartment.[6] Amino acids including arginine, liberated by the catabolism of the proteins, either recirculate or are passed into the bladder. The recirculation includes filtering in the kidney. The bladder store is emptied by excretion of urine.

The urea arrives after synthesis in an initial mixing pool which includes the bloodstream. This pool equilibrates with another pool outside the bloodstream. Urea is lost from the initial mixing pool, either travelling into the gut where it is catabolised, or being filtered by the kidneys and hence passed to the bladder.

Much previous work on urea metabolism in particular, has assumed that the kidneys act either as a mixing pool or mechanical delay. Here an attempt has been made to consider the action of the kidneys in greater detail.

The kidneys consist of an assembly of sub-units called nephrons, each consisting of a blind ended tube into which a tuft of blood capillaries, called a glomerulus, is invaginated (Fig.2). The remainder of the tube is termed the tubule. The glomerulus acts as a filter for plasma, allowing substances of lower molecular weight such as urea and arginine to pass into the tubules, while the proteins remain in the capillaries. After the blood has passed through the glomerulus it enters the efferent arteriole which leads to another capillary network round the tubule. At this stage the tubule cells selectively reabsorb the constituents of the filtrate as they pass along the tubule, preferentially absorbing arginine compared with urea. Once selective reabsorption has occurred giving rise to an interstitial pool of substances, the constituents will either be secreted back into the tubules or secreted into the blood capillaries and hence into the circulation.[7,8,9].

The only other path to be considered is that taken by labelled plasma proteins, other than albumin, synthesised in the liver. In general these too are distributed and are catabolised.

3. MATHEMATICAL MODEL OF THE METABOLIC SYSTEM

3.1 General

The metabolic system described in the previous section thus consists of a series of interconnected pools.

To formulate the mathematical model it is assumed that the rate of change of the quantity of a particular substance in each pool is equal to the summation of two effects. The first is the loss of substance from the pool to other sites and is a function of the quantity of the substance in the pool. The other effect is the gain of the substance from other pools.

The functions concerned are in general non linear. On biological grounds it can be assumed that a labelled substance behaves in the same manner as the same substance unlabelled.

3.2 Model notation

Variables

A - mass of unlabelled albumin.
a - mass of labelled albumin.
C - mass of unlabelled carbon.
c - mass of labelled carbon.
f - general function.
M - mass of unlabelled liver proteins.
m - mass of labelled liver proteins.
Q - fluid mass flow rated (kidney tubules).
q - blood mass flow rates.
R - mass of unlabelled arginine.
r - mass of labelled arginine.
T - delay in albumin pool.
t - time.
U - mass of unlabelled urea.
u - mass of labelled urea.
Z - mass of unlabelled plasma protein other than albumin.
z - mass of labelled protein other than albumin.

All pool sizes are in general functions of time.

E - shift operator. $E^{-1}f(t) = f(t-T)$

Subscripts:- (indicating location of substance)

a - liver albumin synthesis pool.
b - blood.
c - catabolic site of albumin to arginine.
d - liver albumin storage pool.
e - entra-vascular albumin pool.
g - glomerulus, kidneys.
h - arginine catabolic.pool.
i - interstitial pool, kidneys.
j - catabolic site of other proteins to arginine.
k - other protein paths.
l - liver carbon pool.
m - liver cell protein pool.
n - extra-bloodstream urea pool.
o - bladder loss of urine.
p - albumin plasma pool.
q - urea initial mixing pool.
r - liver arginine pool.
s - bladder storage pool.
t - kidney tubules.
u - liver urea pool.
v - catabolic pathways of urea from gut.
w - catabolic site of urea in gut.
x - other pathways for carbon in liver.
y - liver carbon recycling pool.
z - synthesis site of proteins other than albumin.

eg. C_y - mass of unlabelled carbon in liver recycling pool.

U_i - mass of urea in kidney interstitual pool.

Double subscripts have the same meanings as above but indicate transfer from one location to another.

$f_{ap}(\ldots..)$ albumin liver ⟶ plasma

$f_{ti}(\ldots..)$ kidney tubules ⟶ kidney interstitial pool.

3.3 General Rate Equations

3.3.1 Quantity of carbon in the liver carbon pool.

$$\frac{d}{dt}(C_l + c_l) = f_{bl}((C_b + c_b), i_{-l}) + f_{yl}(C_y + c_y) - f_{ly}(C_l + c_l) - f_{lx}(C_l + c_l)$$

3.3.2 Quantity of carbon in liver recycling pool.

$$\frac{d}{dt}(C_y + c_y) = f_{ly}(C_l + c_l) - f_{yl}(C_y + c_y) - f_{yx}(C_y + c_y)$$

3.3.3 Quantity of arginine in liver.

$$\frac{d}{dt}(R_r + r_r) = f_{hr}(C_l + c_l) + f_{hr}(R_h + r_h) + f_{tr}(R_t + r_t) + f_{mr}(M_m + m_m) - f_{ru}(R_r + r_r) - f_{rw}(R_r + r_r) - f_{rz}(R_r + r_r)$$

3.3.4 Quantity of arginine in catabolic pool.

$$\frac{d}{dt}(R_h + r_h) = f_{ch}(A_p + a_p) + f_{jh}(Z_z + z_z) - f_{hg}(R_h + r_h) - f_{hr}(R_h + r_h)$$

3.3.5 Quantity of liver cell proteins in liver.

$$\frac{d}{dt}(M_m + m_m) = f_{rm}(R_r + r_r) - f_{mr}(M_m + m_m)$$

3.3.6 Quantity of arginine in tubules.

$$\frac{d}{dt}(R_t + r_t) = f_{gt}(R_h + r_h) - f_{tr}(R_t + r_t) - f_{ts}(R_t + r_t)$$

$f_{hg} = f_{gt}$ since glomerulus acts as a filter not as a pool.

3.3.7 Quantity of albumin in liver synthesis pool.

$$\frac{d}{dt}(A_a + a_a) = f_{ra}(R_r + r_r) - f_{ad}(A_a + a_a)$$

3.3.8 Quantity of albumin in liver storage pool.

$$\frac{d}{dt}(A_d + a_d) = f_{ad}(A_a + a_a) - f_{dp}(A_d + a_d)$$

3.3.9 Quantity of albumin in plasma.

$$\frac{d}{dt}(A_p + a_p) = E^{-1} f_{dp}(A_d + a_d) + f_{ep}(A_e + a_e) - f_{pe}(A_p + a_p) - f_{pc}(A_p + a_p)$$

3.3.10 Quantity of albumin in extra vascular pool.

$$\frac{d}{dt}(A_e + a_e) = f_{pe}(A_p + a_p) - f_{ep}(A_e + a_e)$$

3.3.11 Quantity of other proteins synthesised.

$$\frac{d}{dt}(Z_z + z_z) = f_{rz}(R_r + r_r) - f_{zj}(Z_z + z_z) - f_{zk}(Z_z + z_z)$$

3.3.12 Quantity of urea in the liver.

$$\frac{d}{dt}(U_u + u_u) = f_{ru}(R_r + r_r) - f_{uq}(U_u + u_u)$$

3.3.13 Quantity of urea in the initial mixing pool.

$$\frac{d}{dt}(U_q + u_q) = f_{uq}(U_u + u_u) + f_{nq}(U_n + u_n) - f_{qn}(U_q + u_q) - f_{qs}(\dot{q}_s, (U_s + u_s), (U_q + u_q)) + f_{sq}(\dot{q}_s, (U_s + u_s), (U_q + u_q)) - f_{qg}(\dot{q}_g, (U_q + u_q)) + f_{iq}(\dot{q}_i, (U_i + u_i), (U_q + u_q)) - f_{qw}(U_q + u_q) + f_{wq}(U_w + u_w)$$

3.3.14 Quantity of urea in the catabolic site (gut).

$$\frac{d}{dt}(U_w + u_w) = f_{qw}(U_q + u_q) - f_{wq}(U_w + u_w) - f_{wu}(U_w + u_w)$$

3.3.15 Quantity of urea in the extra-bloodstream pool.

$$\frac{d}{dt}(U_n + u_n) = f_{qn}(U_q + u_q) - f_{nq}(U_n + u_n)$$

3.3.16 Quantity of urea in tubules.

$$\frac{d}{dt}(U_t + u_t) = f_{qg}(\dot{q}_g, (U_q + u_q)) + f_{it}(\dot{Q}_t, (U_t + u_t), (U_i + u_i)) - f_{ti}(\dot{Q}_t, (U_t + u_t), (U_i + u_i)) - f_{ts}(U_t + u_t)$$

$f_{gt} = f_{qg}$ since the glamerulus acts as a filter and not as a pool.

3.3.17 Quantity of urea in the interstitial pool.

$$\frac{d}{dt}(U_i + u_i) = f_{ti}(\dot{Q}_t, (U_i + u_i), (U_t + u_t)) - f_{it}(\dot{Q}_i, (U_i + u_i), (U_t + u_t)) - f_{iq}(\dot{q}_i, (U_i + u_i), (U_q + u_q))$$

3.3.18 Quantity of urea in the bladder store.

$$\frac{d}{dt}(U_s + u_s) = f_{ts}(U_t + u_t) + f_{qs}(\dot{q}_s, (U_q + u_q), (U_s + u_s)) - f_{sq}(\dot{q}_s, (U_q + u_q), (U_s + u_s)) - f_{sq}(U_s + u_s)$$

4. DYNAMIC TESTS

4.1 Test Method

The experiments carried out all involved the administration of radioactively labelled compound by intravenous injection. Measurements of the concentration of the label in different compounds were made at various sites in the body, yielding impulse responses for sections of the metabolic system.

The experiments were performed on volunteers at the Royal Free Hospital. Data was obtained from ten such patients; some have a normal metabolism and others suffer from diseases of the liver, intestine and other organs.

The experimental method suffers from the limitation that both the quantity of radioactivity injected and quantity and frequency of blood extraction are restricted.

At the present time, five separate experiments are being carried out.

1(a) Sodium Carbonate labelled with ^{14}C is injected into the bloodstream and labelled albumin concentration in the plasma is measured.

(b) A similar test is carried out measuring labelled urea concentration in the mixing pool.

2. Albumin labelled with ^{131}I or ^{125}I is injected into the bloodstream and measurment of the label concentration made in the plasma.

3. Urea labelled with ^{13}C is injected into the bloodstream and the label concentration in the initial mixing pool measured.

4. Sodium carbonate labelled with ^{14}C is injected into the blood and the label concentration in blood carbonate- bicarbonate measured.

In order to clarify the model of the metabolism, further experimental data would be required.

4.2 Test Results

The experiments are typified by the curves shown in Fig.5.

All the results exhibited considerable scatter, this being most pronounced in the albumin data. This is attributable to the effect of background radiation, the low levels of radiation involved and fluctuations in the bodily functions.

5. MODEL REDUCTION

With the very restricted methods of dynamic testing which can be applied it is of course not possible to identify fully the model described in Section 3.

It was therefore necessary to reduce this model. Extensive work has recently been carried out into methods of model reduction.[10,11,12] In this case, it was decided in the first instance to adopt the simplest approach. The general rate equations were linearised and further reductions carried out on the linearised model to remove the less significant modes. The resultant model is not rigorously accurate but it is hoped that it will give useful biological information. It is hoped in the future to improve the technique as understanding of the biology improves.

5.1 Linearisation of rate equations.

In the evaluation of the tests we are only concerned with the labelled substances. Since the quantity of labelled substance in any pool is small compared with the quantity of unlabelled substance and since the latter can be assumed to remain constant, the general rate equations can be linearised. Thus equation 3.3.1 reduces to 5.1.1.

5.1.1 Carbon in liver carbon pool.

$$\frac{dc_l}{dt} = \frac{\partial f_{lu}}{\partial C_b} \cdot c_b + \frac{\partial f_{lu}}{\partial \dot{q}_l} \cdot \dot{q}_l + \frac{\partial f_{ly}}{\partial C_y} \cdot c_y - \frac{\partial f_{ly}}{\partial C_l} \cdot c_l - \frac{\partial f_{lr}}{\partial C_l} \cdot c_l$$

3.3.2 to 3.3.18 then are of similar form.

After investigating the further treatment of these equations, by time domain matrix methods it was decided that considerable simplification would

result by putting these equations into transfer function form and that this would not introduce any difficulties.

The transfer function block diagrams representing the linearised rate equations are shown in Fig.3 and 4.

5.2 Dynamic reduction

5.2.0 Principles.

The linearised rate relations in transfer function form were then examined. From biological information a table was prepared of the time constants associated with the various pools.

Pool	Time-Constant(Hours)	Pool	Time-Constant(Hours)
Liver Carbon	10^{-3}	Argenine Tubules	10^{-2}
Liver Recycling[13]	10^{-2}	"Other Protein"[17]	10^{2}
Liver Argenine[14]	10^{-3}	Liver Urea	10^{-3}
Liver Cell Protein	10^{2}-10^{3}	Urea initial mixing pool[13]	10^{-1}
Liver Albumin Synthesis[15]	10^{-3}	Extra bloodstream pool[13]	10^{0}
Liver Albumin Storage[16]	10^{-1}	Gut catabolic pool[18]	10^{0}
Plasma Albumin Pool	10^{-1}	Kidney Interstitial[9,19]	10^{-2}
Extra-vascular Albumin[2]	10^{0}	Urea Tubules[9]	10^{-2}
Catabolic Pool[1]	10^{2}	Bladder Store[19]	10^{-2}

Terms which involved time constants of the order of 10^{-2} hours or less were reduced to pure gains. In addition pathways wherein very small quantities of substance were being fed along paths with time constants of the order of 10^{2} hours or more were also neglected since they would have had little effect over the period of observation.

The dynamics of the liver carbon, arginine, albumin synthesis and urea synthesis pools were thus neglected as fast. The effect of the paths of arginine from the catabolic to the liver pool and of other plasma proteins from the liver arginine to the catabolic pool were taken as negligible. Similarly, feedback of urea from the bladder and the interstitial pool were neglected as small.

5.3 Simplified model

The linearised simplified model is shown in Fig.6. It has been reduced to those gains and time constants which can be identified from the experimental results.

6. ANALYSIS OF EXPERIMENTAL RESULTS

The experimental results of dynamic tests were compared with the simplified model by comparing predicted and actual responses.

A number of methods of fitting experimental responses to predicted response equations were considered.[2,20,21,22] It was decided to adopt a

simple adaptive fitting technique minimising total square error. The fitting was carried out on a digital computer using Rosenbrock's hill climbing technique,[23] and confining the search to the neighbourhood of initial parameter values estimated graphically.[2]

The dynamic responses for three patients have been analysed in this way.

The responses are found to fit the predictions of the simplified model with an accuracy of the order of 4% (r.m.s. error of observation), which is the same order as experimental accuracy.

7. CONCLUSIONS

A mathematical model of albumin and urea metabolism consistent with current biological concepts has been derived.

In carrying out this interdisciplinary investigation a fresh look had to be taken at some of the biological problems involved and a clearer comprehension of the process has been obtained.

It is considered from this experience that the application of dynamic analysis methods to pool dynamics problems can be generally helpful in clarifying concepts and assumptions.

The investigation is still in progress. Further test data are being analysed and other response fitting methods are being investigated. The future programme envisages, as mentioned before, additional biological tests and detailed analysis of the biological significance of the estimated parameter values. The model obtained must thus be viewed only as a first approach to the truth.

8. ACKNOWLEDGMENTS

We wish to thank in particular Dr. V.M. Rosenoer and Dr. A. Craigie of the Royal Free Hospital for their great help in this work. It was in particular due to Dr. Rosenoer's interest that this investigation was initiated. Two of us (C.D.B. and E.R.C.) were in receipt of grants from the Science Research Council while this work was carried out and would like to record their thanks.

REFERENCES

(1) Cohen S., Freeman T. & Macfarlane A.S., Clinical Science 1961, 20, 161.

(2) Matthews C.M., Physics in Medicine and Biology 1957, 2, 36.

(3) Berman M., Shahn E., Weiss M.F., Biophys. Journal 1962, 2, 289.

(4) Reeve E. & Bailey, J. Lab. Clinical Medicine 1962, 60(6), 923

(5) Swick R.W., J.Biol.Chem. 1958, 231, 751.

(6) Beeken, W.L. et al, J.Clin.Invest. 1962, 41, 1312.

(7) Green, "Introduction to Human Physiology", Oxford University Press 1965.

(8) Keele & Neil, "Sampson Wright's Applied Physiology", OUP, 1965.

(9) Robinson J.R., "Renal Disease", ed. Black, ch.3, Blackwell, 1967.

(10) Anderson J.H., Proceedings IEE 1967, 114(7), 1014.

(11) Davison E.J., IEEE Transactions on Automatic Control, 1966, 11, 93.

(12) Mitra D., AEEW-R520, 1967.

(13) Jones E.A., et al, Clinical Science 1968 (in press).

(14) Delluva A.M. & Wilson D.W., J.Biol.Chem. 1946, 166, 739.

(15) Craddock V.M. & Dalgleish C.E., Biochem. J., 1957, 66, 250

(16) Peters T., J.Biol.Chem., 1962, 237, 1186.

(17) Schultze H.E. & Heremans J.F.(ed), "Molecular biology of human proteins", Vol.1, Section 3, Ch.2, (Elsevier Publishing Co.)

(18) Waber M. & Bodenlos L.J., J.Clin.Investigation, 1958, 38, 1617.

(19) Regeoczi E., Irons L., Koj A., Macfarlane A.S., Biochem.J. 1965, 95, 521.

(20) Hildebrand F.B., "Introduction to Numerical Analysis" (McGraw Hill 1956)

(21) Hudson G.E., American J.Physics, 1953, 21, 362.

(22) Householder A.S., U.S. Atomic Energy Commission Report ORNL-455(1950).

(23) Rosenbrock H.H., & Storey C., "Computational Methods for Chemical Engineers" (Pergamon 1966).

^{14}C PATHWAYS

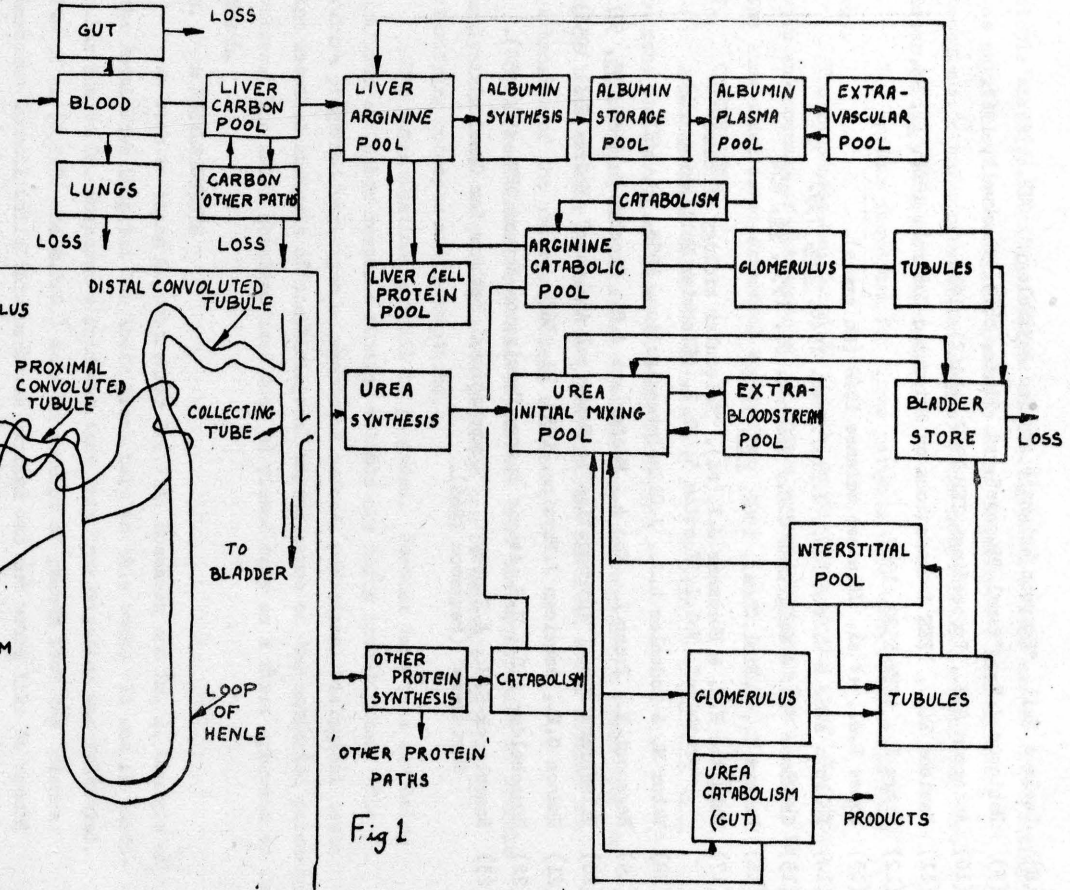
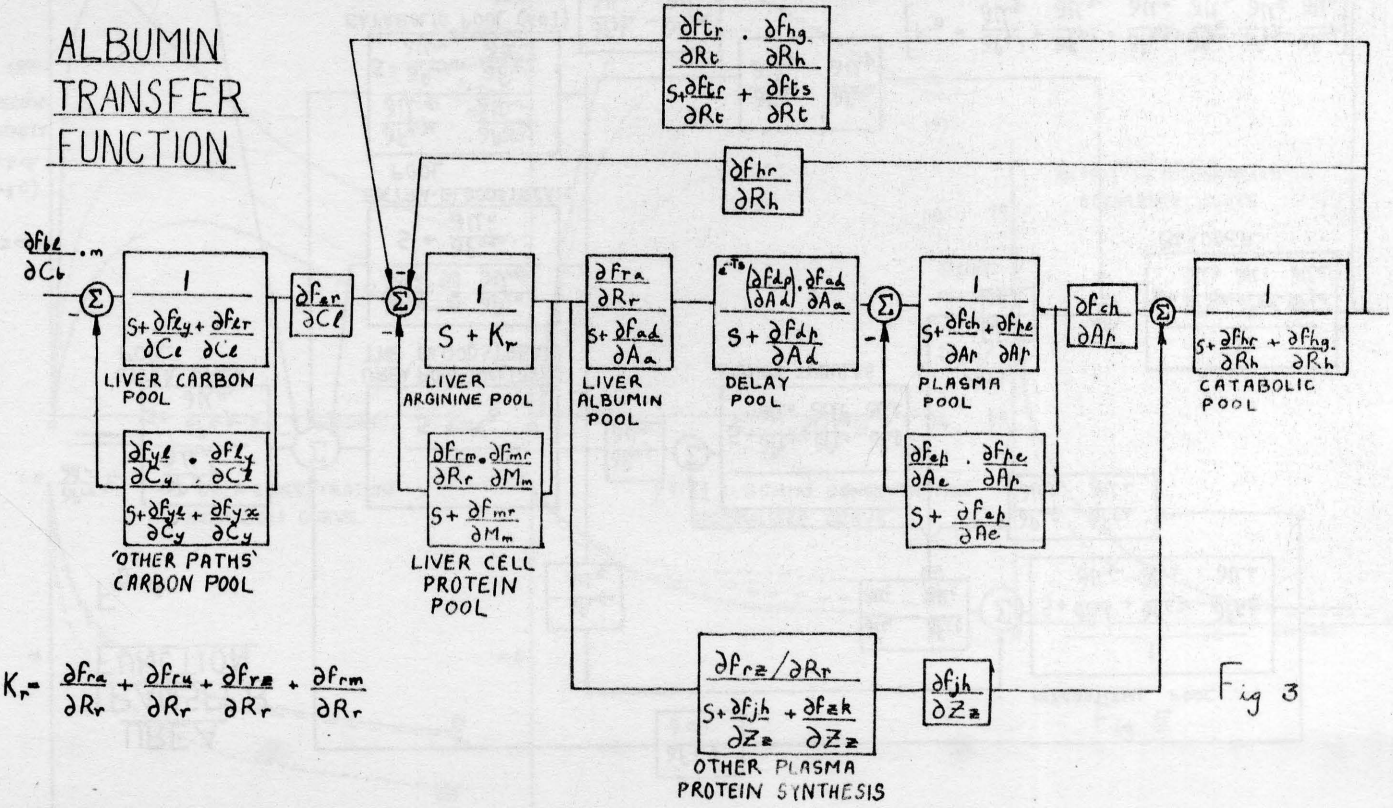


Fig 1

A TYPICAL KIDNEY NEPHRON AND GLOMERULUS

Fig 2

ALBUMIN TRANSFER FUNCTION

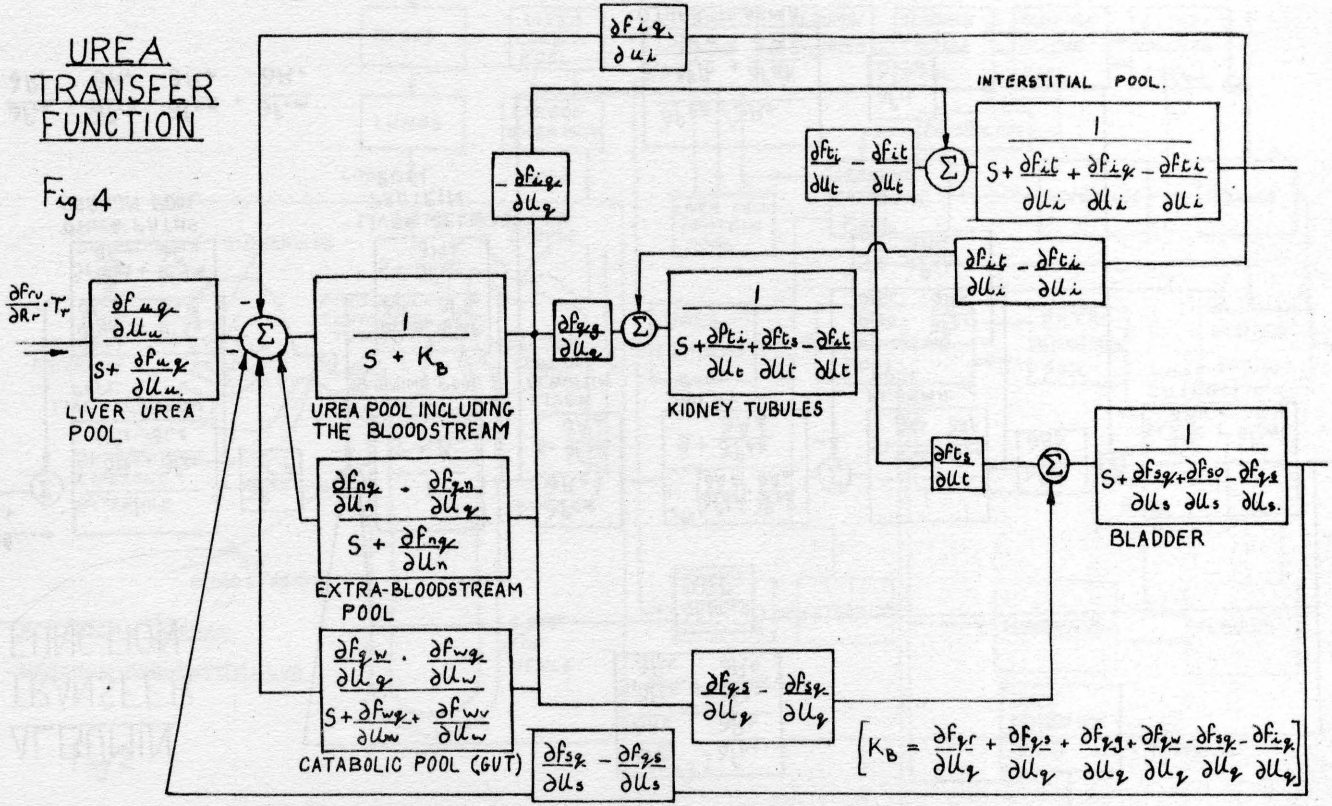


$$K_r = \frac{\partial f_{ra}}{\partial R_r} + \frac{\partial f_{r4}}{\partial R_r} + \frac{\partial f_{rz}}{\partial R_r} + \frac{\partial f_{rm}}{\partial R_r}$$

Fig 3

UREA TRANSFER FUNCTION

Fig 4



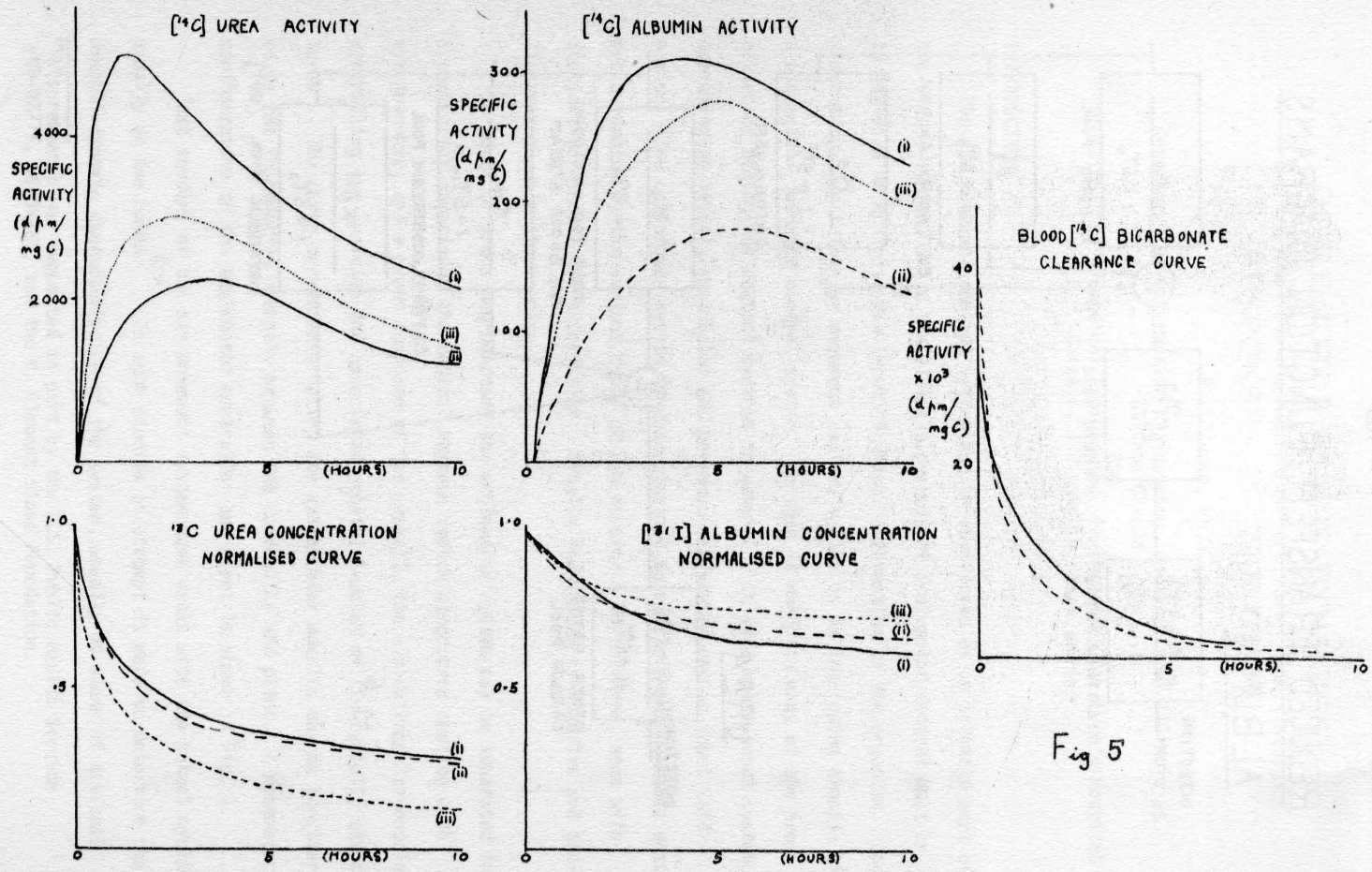
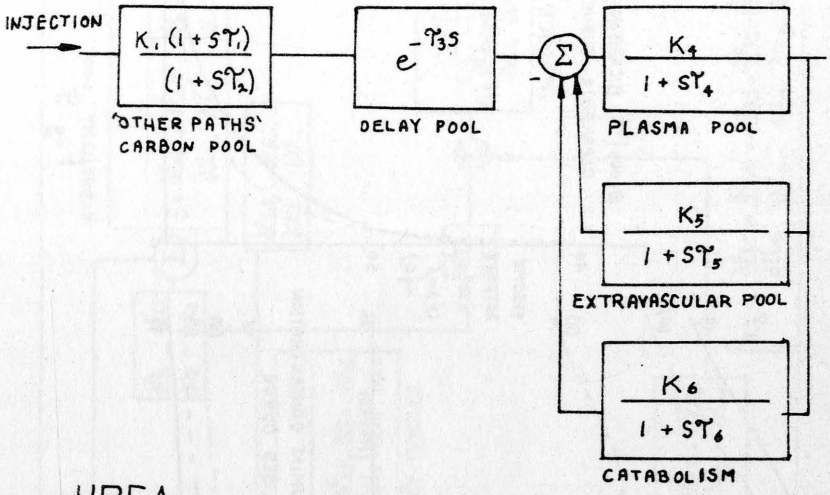

[14C] UREA ACTIVITY
4000
SPECIFIC ACTIVITY (dpm/mg C)
2000
(i)
(iii)
(ii)
0
5
(HOURS)
10
[14C] ALBUMIN ACTIVITY
300
SPECIFIC ACTIVITY (dpm/mg C)
200
100
(i)
(iii)
(ii)
0
5
(HOURS)
10
BLOOD [14C] BICARBONATE CLEARANCE CURVE
40
SPECIFIC ACTIVITY x 10³ (dpm/mg C)
20
0
5
(HOURS).
10
13C UREA CONCENTRATION NORMALISED CURVE
1·0
·5
(i)
(ii)
(iii)
0
5
(HOURS)
10
[131I] ALBUMIN CONCENTRATION NORMALISED CURVE
1·0
0·5
(iii)
(ii)
(i)
0
5
(HOURS)
10

Fig 5

REDUCED TRANSFER FUNCTION DIAGRAMS

ALBUMIN



UREA

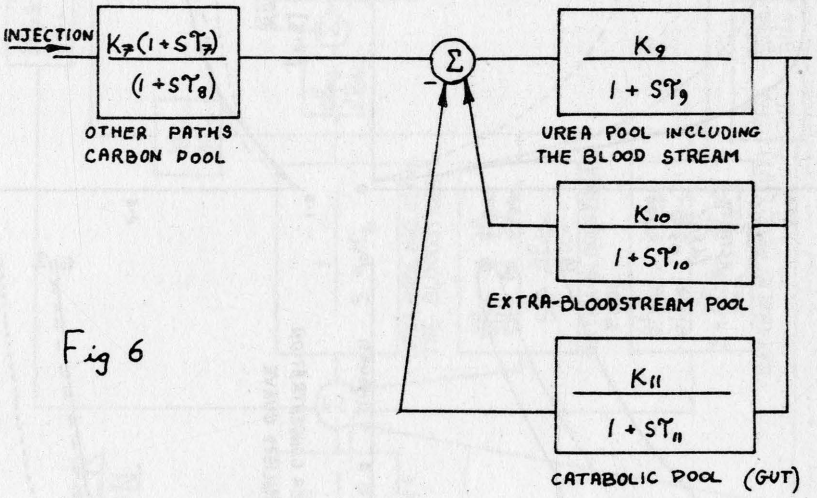


Fig 6

OCULOMOTOR PLANT DYNAMICS: ELECTROMYOGRAPHIC AND TRANSIENT RESPONSES IN THE CAT*

B.L. Zuber, Ph.D.

Department of Information Engineering, University of Illinois at Chicago Circle
and
Department of Biomedical Engineering, Presbyterian-St. Luke's Hospital
Chicago, Illinois, U.S.A.

INTRODUCTION

The eye movement control system can be considered as an intimate part of the visual system, since the ultimate purpose of oculomotor control must be to maximize or optimize the input of visual information to the organism. Considered by itself the eye movement control system is an intriguing example of a biological feedback control system. As such it exhibits many of the properties of biological control systems in general. It is a multi-input system, operating on visual, vestibular and perhaps neck proprioceptor input information. The oculomotor system is characterized by multiple output modes which are evidenced by the various types of eye movements it produces, each with distinctive dynamic characteristics. Finally the system is adaptive and nonlinear and non-minimum phase.

In terms of gross neuroanatomy the oculomotor system may be separated into a controller comprising the central nervous system structures dealing with eye movement; and a plant composed of the eyeball, the six extraocular muscles controlling eye position and associated orbital (socket of the eyeball) structures[1]. The three nerves innervating the extraocular muscles may be considered as the information channels between the controller and plant. A schematic configuration of the oculomotor controller and plant is shown in Fig. 1.

Past studies of the eye movement system have dealt with the intact system, usually in the human[2-9]. In such studies an attempt is made to determine the overall dynamic characteristics of the system, usually by means of a visual

*This research was supported in part by the U.S. Public Health Service (NB-07777, NB-06487) and the W. Clement Stone Foundation.

tracking task. The position of the tracking target is considered as the system input and the position of the eye as the system output. Such studies are extremely valuable not only for the dynamic characteristics they uncover, but also because they are performed on an intact normally operating biological system (i.e., no dissection). These studies do not, however, allow specific dynamic elements to be unequivocally localized or associated with specific anatomical structures in the controller or plant.

More recent studies of oculomotor control have emphasized interactions between the controller and plant and have attempted to determine the plant dynamics as they are related to structural entities within the orbit[1,10,11]. These experiments have involved detailed studies of eye movement trajectories[1,12], determination of active and passive properties of the extraocular muscles under isometric and isotonic conditions in humans and experimental animals[10], and extrapolation of physiological data from the literature[1]. An interesting and valuable contribution of these studies has been and indication of the form of the controller output associated with the saccadic eye movement[1,10]. The saccadic eye movement is the basic step response of the visual fixation reflex and is used for shifting fixation from one point in visual space to another. This type of movement is characterized by very high velocity and exhibits interesting amplitude-dependent non-linear properties[13]. All studies of the oculomotor plant dynamics indicate that the plant dynamics are too slow to account for the dynamics represented by the saccade[1,10,14]. Therefore, it is concluded that the slow plant dynamics are overcome by using a non-stepwise controller output (plant input) to produce the saccadic eye movement[1,10].

Because all types of eye movements are produced by the final common pathways of the oculomotor system (ocular motor nerves and extraocular muscles) their trajectories must reflect the dynamics of the oculomotor plant. This means that any model of the plant must be capable of accounting for a wide

variety of oculomotor responses, including the saccadic eye movement.

A relatively direct approach to the determination of the plant dynamics involves isolation of the plant and gaining access to one or more of its input channels, the motor nerves serving the extraocular muscles[14,15]. This paper describes experiments based on this approach, involving breaking into the oculomotor "black box" of the anesthetized cat. Frequency modulated pulse trains have been injected into the oculomotor nerve to cause eye movement due to changes in the contractile state of the medial rectus muscle. Experiments have been designed to answer basic questions regarding system linearity and the degree to which events occurring in the muscle during stimulation resemble those which occur during muscle contractions under more physiological conditions. Frequency response experiments have been performed using a sinusoidal input while measuring an intermediate output, the electromyogram. Finally, eye movements resembling saccadic eye movements have been produced by assuming the form of the plant input signal suggested by previous studies[1,10].

METHODS

The Preparation. Cats were anesthetized with 30 mg/kg of sodium pentobarbital administered intraperitoneally. The femoral vein was catheterized and anesthesia was maintained by intravenous administration of sodium pentobarbital at a level where corneal and lid reflexes were absent. The animal was aligned in a stereotaxic frame, the scalp incised and retracted, and craniotomy was performed to allow insertion of the stimulating electrode. Pressure points and incisions were covered with viscous xylocaine.

The Stimulus. Pulse train stimuli were injected into the oculomotor nerve within the brain stem by means of a sterotaxically placed concentric bipolar electrode. Pulses were obtained from a physiological stimulator with isolated output. This stimulator was triggered by a voltage-controlled pulse generator whose output pulse frequency was proportional to the voltage level at the voltage-control terminals of the pulse generator. Thus, a stimulus

pulse train could be generated with pulse frequency proportional to the voltage level of any modulating waveform selected by the experimenter. Modulating waveforms used in the experiments described below consisted of sinusoids, steps and pulse-step combinations.

Recording. Eye movements were recorded using a modification of a previously described differential reflection technique[6,16]. In the present case infrared-sensitive light sensors placed in a horizontal plane at the level of the center of the pupil, and arranged symmetrically on either side of the pupil. The light sensors were arranged in a bridge-amplification circuit which gave a voltage output proportional to eye position. The pupil of the measured eye was constricted with topical application of physostigmine. Eye velocity was obtained by analog differentiation of the eye position signal.

The electromyogram (EMG) of the medial rectus muscle was recorded by means of a 0.003 inch teflon-coated wire inserted into the muscle from the front of the orbit. Insulation was stripped off at the very end to provide a recording surface. The EMG was amplified with a low-level AC coupled preamplifier with a low frequency cut off point of 8 Hz and a high frequency cut off point of 1 kHz.

RESULTS

The experiment began by driving the stimulating electrode through the cerebrum toward the location of the third nerve within the brain stem as selected from the sterotaxic atlas[17]. As soon as the brain stem was entered the stimulus was applied for 2-5 sec after each 1 mm increase in depth of the electrode. The stimulus took the form of either a train with pulse frequency sinusoidally modulated between 50 and 150 pps, or a constant pulse frequency train. The physiological effect of the stimulus was observed. At a point 5-7 mm above the target location the first effects of stimulation appeared. These consisted of bilateral pupillary dilatation and an increase of unsynchronized EMG activity of the ipsilateral medial rectus. At this point

penetration, and subsequent stimulation, was continued in 0.5 mm steps. By careful manipulation of electrode depth a position could be found near the target at which large stimulus synchronized muscle potentials appear on the EMG trace and where horizontal eye movement was maximal. Pupillary dilation was not observed. The electrode was left in this position and experimental runs commenced. At this point, depending on the nature of the experiment, the modulation waveform controlling stimulus pulse frequency was selected by the experiments.

Controls. It was important to show, in so far as possible, that the medial rectus was the only extraocular muscle contracting in response to stimulation and that this muscle was contracting as a result of action potentials transmitted to it over its motor nerve, the oculomotor nerve. The first step in this process was to determine the location of the stimulating electrode. After most experiments the animals were sacrificed, the brains were removed and fixed in formalin. Ten micron serial sections of the brain stem were stained and mounted to aid in localizing the electrode tract. Fig. 2 is a transverse section of the brain from one animal showing the tip of the electrode tract (broken oval on right hand side). The entire tract does not appear because the section is not truly parallel to the tract. On the opposite side of the brain the oculomotor nerve may be seen emerging from the brain stem (smaller circle). The actual point of stimulation is about 1 mm above the bottom of the brain stem as shown by the arrow. At this point the fibers of the oculomotor nerve are collecting to form the complete nerve in preparation for exit from the brain stem. In some sections they may be seen to have collected into discrete bundles within the brain stem. Fig. 2 confirms the location of the stimulating electrode to be among oculomotor nerve fibers.

Since the oculomotor nerve innervates four of the six extraocular muscles it is important to show that the medial rectus is the only muscle responding to the stimulus. Gross observation of the eye movements produced by

stimulating indicated that this was the case. The direction of eye movement was always medial, and the higher the stimulus frequency the more medial the position of the eye with respect to the straight-ahead position (see below). In one cat, after the stimulating electrode was in place, the ipsilateral eye was collapsed and removed after attaching sutures to all six extraocular muscle insertions. Observation of all six muscles indicated that only the medial rectus contracted in response to stimulation. The other five muscles were apparently unaffected by the stimulus.

Finally, to investigate the function of the neuromuscular junction during stimulation, the effect of neuromuscular blocking agents was studied. Both decamethonium bromide and gallium arsenide were used to block neuromuscular transmission between the oculomotor nerve and the medial rectus muscle. Fig. 3 shown the effect of injecting 0.5 mgm of decamethonium bromide into the medial side of the orbit. The figure shows the changes which occur in the muscle potentials produced by a single pulse stimulus at various times after the injection of the drug. The multiphasic waveforms shown in Fig. 3, while they have been observed in several experiments, are not representative of the muscle potentials usually recorded from this preparation. The more representative waveform is shown in Fig. 5a. After 5 minutes the muscle potential was no longer detectable, then it gradually increased in amplitude until 62 minutes after injection, when it had almost returned to the pre-injection amplitude. The results with gallium arsenide were similar to those shown in Fig. 3.

Linearity and Static Calibration. In the next series of experiments it was desired to determine the input-output characteristics of the system with stimulation pulse frequency as the input and eye position as the output. The static calibration curve of Fig. 4 shows the results of one such experiment and represents the input-output characteristic of the system. In this experiment the steady state position of the front of the eye was measured using a

microscope with a calibrated ocular while the nerve was stimulated with a constant pulse frequency. The ordinate in Fig. 4 represents eye position with respect to the primary (straight ahead) position. Movement of the front of the eye was converted to ocular rotation in degrees by simple triangulation. Fig. 4 indicates that at stimulation frequencies at and above 200 pps the system was operating in a non-linear range of saturation. However, a very nearly linear operating range existed for stimulus pulse frequencies between 50 and 150 pps. This corresponded to a range of eye movement of about 12 degrees. Within this range the system was characterized by a sensitivity of about 0.12 degrees/pps[18].

Frequency Response. Within the linear operating range of the system the use of sinusoidally modulated stimulus pulse frequencies produced sinusoidal eye movements. For modulating waveform frequencies up to about 25 Hz the system may be characterized as linear first order with a time constant of about 64 msec. The system contains a non-minimum phase element equivalent to a pure delay of 12 to 15 msec. Since frequency response experiments have been previously published[14] they will not be reported in detail here.

Electromyogram. The electromyogram represents an intermediate process within the system. It presumably represents the muscle potential propagated along the muscle fiber as a result of activation of the neuromuscular junction. Since the muscle potentials recorded from this preparation were synchronized with the stimulus pulses, it was relatively easy to determine the dynamics of the process or processes underlying the generation of the muscle potentials.

Fig. 5a shows the typical form of the muscle potentials recorded from this preparation. The figure shows superimposed traces of muscle potentials produced by single pulse stimuli of approximately 0.1 msec duration. The delay between the onset of the stimulus pulse and the onset of the response was about 1.0 msec as might be expected from a single neuromuscular junction delay. Fig. 5b shows the EMG response to constant pulse frequency stimulus

trains ranging from 100 to 1000 pps. The interval between the first two muscle potentials in each train was aberrant and should be ignored. However, in all cases the first response was the largest of the train, indicating that all subsequent pulse intervals were so short that responses interacted, as evidenced by the reduction in amplitude of the train as compared with the first muscle potential. In this experiment the input was again stimulus pulse frequency, but the output was now considered as frequency of occurrence of EMG potentials. Observation of the steady state portions the responses in Fig. 5b (ignoring amplitude changes) indicates that the input-output curve for this subsystem (oculomotor nerve - EMG) is linear with a slope of +1 (EMG frequency equals stimulus pulse frequency) up to a stimulus pulse frequency of about 700 pps. At 1000 pps some periodic responses were observed although this frequency must very nearly correspond to the absolute refractory period of the nerve.

In order to determine the contribution of this subsystem to the dynamics of the overall system (oculomotor nerve-eye position) sinusoidally modulated pulse trains were used. This input corresponded exactly to that used in the frequency response experiments mentioned above. Stimulus pulse frequency was sinusoidally modulated between 50 and 150 pps, with modulation waveform frequencies ranging from 0.5 to 20 Hz. EMG potentials triggered standard pulses were integrated using an analog integrator which provided a DC voltage proportional to average frequency. The dynamics of the integrator were determined by using the stimulus train as an input to the integrator after each experimental run. These dynamics were subtracted from the recorded response dynamics to derive the EMG subsystem dynamics. Some recorded responses and modulating waveform from one experiment are shown in Fig. 6. Cycle-by-cycle average responses were taken on-line by a digital computer (IBM 1800) which also performed Fourier analysis on responses to provide gain and phase data for frequency response computations. The computer plotted one cycle of the average response and modulation waveform for each run at a given modulation waveform frequency. Notice that the

responses are sinusoidal, again indicating linearity. The attenuation of the average response waveform and its phase shift with respect to the modulating waveform as shown in Fig. 6 actually represent the dynamics of the integrator. Once these dynamics were subtracted out the responses showed virtually no dynamics with the exception of a small phase lag which increased linearly with frequency, reflecting the small neuromuscular junction delay mentioned above. The adjusted gain points scattered randomly about unity gain. These dynamics are characteristic of a non-minimum phase element, in this case the pure delay represented by the neuromuscular junction delay of about 1.0 msec.

Transient Eye Movement Responses. The use of transient inputs allowed a confirmation of frequency response data obtained with sinusoidal inputs, and also provided a direct visualization of the effect of the non-minimum phase element (12-15 msec delay) in the time domain. A step change in stimulus pulse frequency from 50 to 150 pps resulted in a step change in eye position with a time constant of about 64 msec as predicted by the frequency response data. The non-minimum phase element manifested itself as an apparent delay of 12 to 15 msec between onset of stimulation and onset of eye movement. This again confirms the observation of the non-minimum phase element indicated in the frequency response. The delay in the step response could not be accounted for by a delay between the onset of the stepwise modulating waveform and the onset of the actual change in stimulus pulse frequency. Further support for the idea that the response delay was not produced by an artifact was seen in the control experiments described under the EMG results. When the sinusoidally modulated pulse train stimulus was integrated and the integral Fourier analyzed the resulting frequency response was that predicted by the integrator circuitry. A delay between modulating waveform and stimulus pulse train would have shown up in this frequency response.

Eye movements resembling saccades have been produced by using as an input a combined pulse-step modulating waveform. Basically this input consists

of an initial high frequency burst of stimulus pulses followed by a lower steady pulse frequency train. The duration of the burst was chosen to correspond to the approximate duration of human saccadic eye movements. The pulse frequency was chosen for minimum overshoot on the resulting eye movement response. The pulse frequency of the steady train determined the final position of the eye, or the amplitude of the "saccadic" eye movement. Fig. 7 shows the eye movement resulting from this type of stimulus. The top trace shows superimposed traces of the eye movement, the middle trace shows the associated eye velocity traces and the bottom trace is the modulating waveform representing stimulus pulse frequency. The oscilloscope was triggered at the onset of the stimulus. The 12 to 15 msec delay is clearly evident in both eye position and velocity traces in this figure. Notice that the eye movements were completed in about forty milliseconds. This figure demonstrates that by manipulating the plant input a basically stepwise response, resembling a saccadic eye movement, can be completed in a time shorter than the basic time constant governing the plant dynamics.

DISCUSSION

The usefulness of a technique such as injecting signals into neurological pathways to partially "dissect" biological systems is considerably enhanced if the system behaves linearly, and if the desired information channels can be specifically stimulated. In some situations the nerves of interest can be dissected out in the periphery and the stimulus applied more directly[19]. In the present case the results of various experiments all strongly suggest that the part of the oculomotor nerve innervating the medial rectus may be specifically stimulated within the brain stem, and that the system has an adequate linear operating range within which a dynamic description of the system may be obtained.

In terms of the electromyographic response of the muscle the experiments would clearly indicate that the muscle is receiving its activation normally.

Of particular interest in this connection is the pharmacological block, the normal neuromuscular junction delay and the apparent refractoriness observed in the experiments. The muscle potentials themselves probably reflect synchronous summated potentials recorded from a number of adjacent motor units. It may be that a suprathreshold stimulus such as that used in these experiments causes all motor units to fire synchronously. The extent of synchrony or recruitment cannot be determined by using a single electrode.

The system having the stimulus to the oculomotor nerve as input and EMG of the medial rectus as output does not appear to contribute to the minimum phase dynamics of the overall system. There appear to be two non-minimum phase elements in the system, the 1.0 msec delay of the neuromuscular junction and the longer delay of 12 to 15 msec between stimulus and eye movement. The origin of this latter delay is not clear at present. The results of the EMG experiments indicate that the minimum phase first order dynamics previously observed are confined to that part of the plant which is beyond the process or processes generating the EMG. These first order plant dynamics are undoubtedly a reflection of the fact that the inertia of the eyeball is essentially negligible[1,10,11].

A first order description of the eye movement plant is considerably simpler than previous proposed models which have been generally higher order and in some cases non-linear[1,10,11]. While these models have proven adequate, it is clear from the above results with transient inputs that a first order description may encompass a wide range of oculomotor activity[20]. In particular, using a combination pulse-step change in stimulus frequently results in what is essentially a saccadic eye movement. Thus, inherently slow plant dynamics may be overcome by using an initial burst of stimulus frequency as predicted by earlier studies[1,10]. The above results are the first actual demonstration that such a predicted input will actually produce the saccadic output. Furthermore, in a digital simulation of the plant dynamics it has been observed

that small changes in the parameters describing the pulse-step input significantly affect the shape of the "saccadic" output[20]. This may suggest that small statistical variation in the controller output may be responsible for variations (e.g. amount of overshoot) in saccadic eye movements.

REFERENCES

1. Cook, G. and L. Stark (1967). Derivation of a Model for the Human Eye-Positioning Mechanism, Bull. Math. Biophys., 29, 153-175.

2. Yarbus, A. L. (1957). Motion of the Eye on Interchanging Fixation Points at Rest in Space, Biofizika, 2, 698-702.

3. Fender, D. H. and P. W. Nye (1961). An Investigation of the Mechanisms of Eye Movement Control, Kybernetik., 1, 81-88.

4. Gurevich, B. K. H. (1961). Universal Characteristics of Fixation Reflexes of the Eye, Biofizika, 6, 377-383.

5. Rashbass, C. and G. Westheimer (1961). Disjunctive Eye Movements, J. Physiol., 159, 339-360.

6. Stark, L., G. Vossius, and L. R. Young (1962). Predictive Control of Eye Tracking Movement. IRE Trans. Human Factors Electron, HFE-3, 52-57.

7. Young, L. R. and L. Stark (1963). A Discrete Model for Eye Tracking Movement. IEEE Trans. MIL. Elec., MIL.-7: 113-115.

8. Dallos, P. J. and R. W. Jones (1963). Learning Behavior of the Eye Fixation Control System, IEEE Trans. Auto Con., AC-8, 218-227.

9. Zuber, B. L. and L. Stark (1968). Dynamical Characteristics of the Fusional Vergence Eye Movement System, IEEE Trans. Sys. Sci. Cyber., SSC-4, 72-79.

10. Robinson, D. A. (1964). The Mechanics of Human Saccadic Eye Movement, J. Physiol. London, 174: 245-264.

11. Childress, D. S. and R. W. Jones (1967). Mechanics of Horizontal Movement of the Human Eye, J. Physiol., 188, 273-284.

12. Cook, G., L. Stark, and B. L. Zuber (1966). Horizontal Eye Movements studied with the On-Line Computer, Arch. Ophthal., 76, 589-595.

13. Zuber, B. L., Semmlow, J. L. and L. Stark (1968). Frequency Characteristics of the saccadic Eye Movement. Biophysical Journal (In press).

14. Zuber, B. L. (1968). Eye Movement Dynamics in the Cat: The Final Motor Pathway, Expl. Neurol. 20, 255-260.

15. Zuber, B. L., Reinhart, R. J. and J. L. Beug (1967). The Final Output Pathways for Oculomotor Control in the Cat, Proc. 20th Ann. Conf. Engineering in Medicine and Biol, 9, 7.2.

16. Zuber, B. L. (1965). Physiological Control of Eye Movements in Humans, PH.D. Dissertation, M.I.T.

17. Snider, R. S. and W. T. Niemer (1961). "A Stereotaxic Atlas of the Cat Brain" Univ. of Chicago Press, Chicago.

18. Zuber, B. L. (1968). Sinusoidal Eye Movements From Brain Stem Stimulation in the Cat, Vision Research, 8, 1073-1079.

19. Partridge, L. D. (1966). Signal-handling Characteristics of Load-moving Skeletal Muscles, Am J. Physiol., 210, 1178-1191.

20. Reinhart, R. J. and B. L. Zuber (1968). Control of Eye Position by the Lateral Rectus Muscle in the Cat, Proc. 21st Ann. Conf. Engineering in Medicine and Biology (In Press).

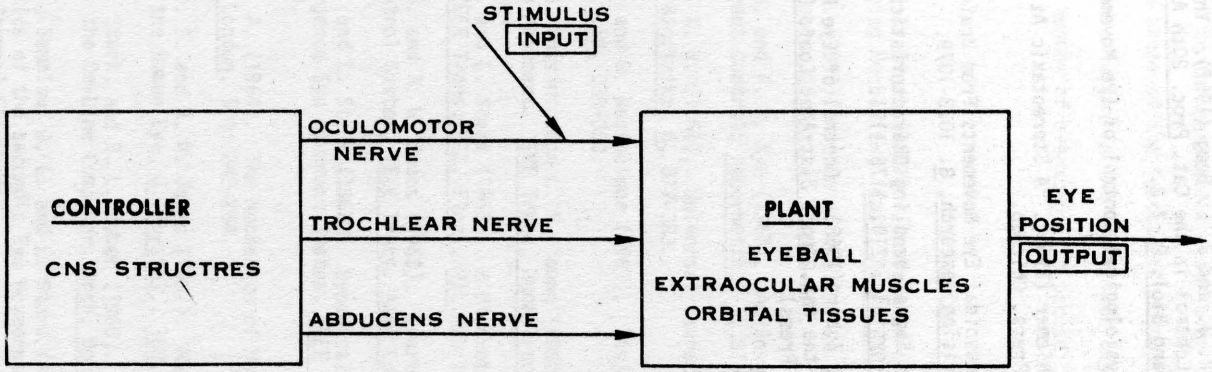


Fig. 1 Schematic division of oculomotor system into controller and plant based on gross anatomy. In the experiments described the input signal is injected into one of the information channels, the oculomotor nerve. Output shown is eye position, but it may be an intermediate output, the electromyogram.

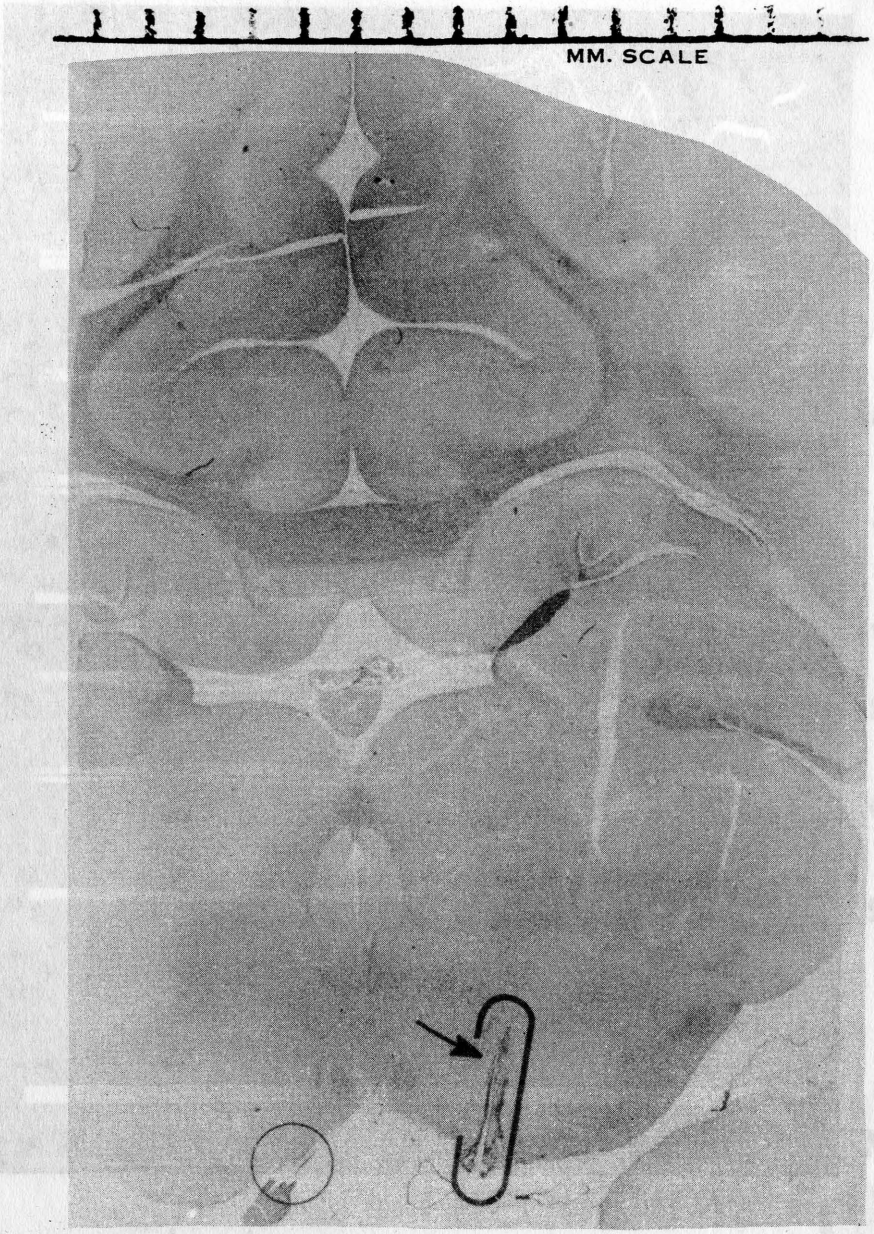


Fig. 2 Photomicrograph showing part of stimulating electrode tract (broken oval). Arrow shows approximate position of active bipolar electrodes. Contralateral oculomotor nerve is seen emerging from brain stem (circle).

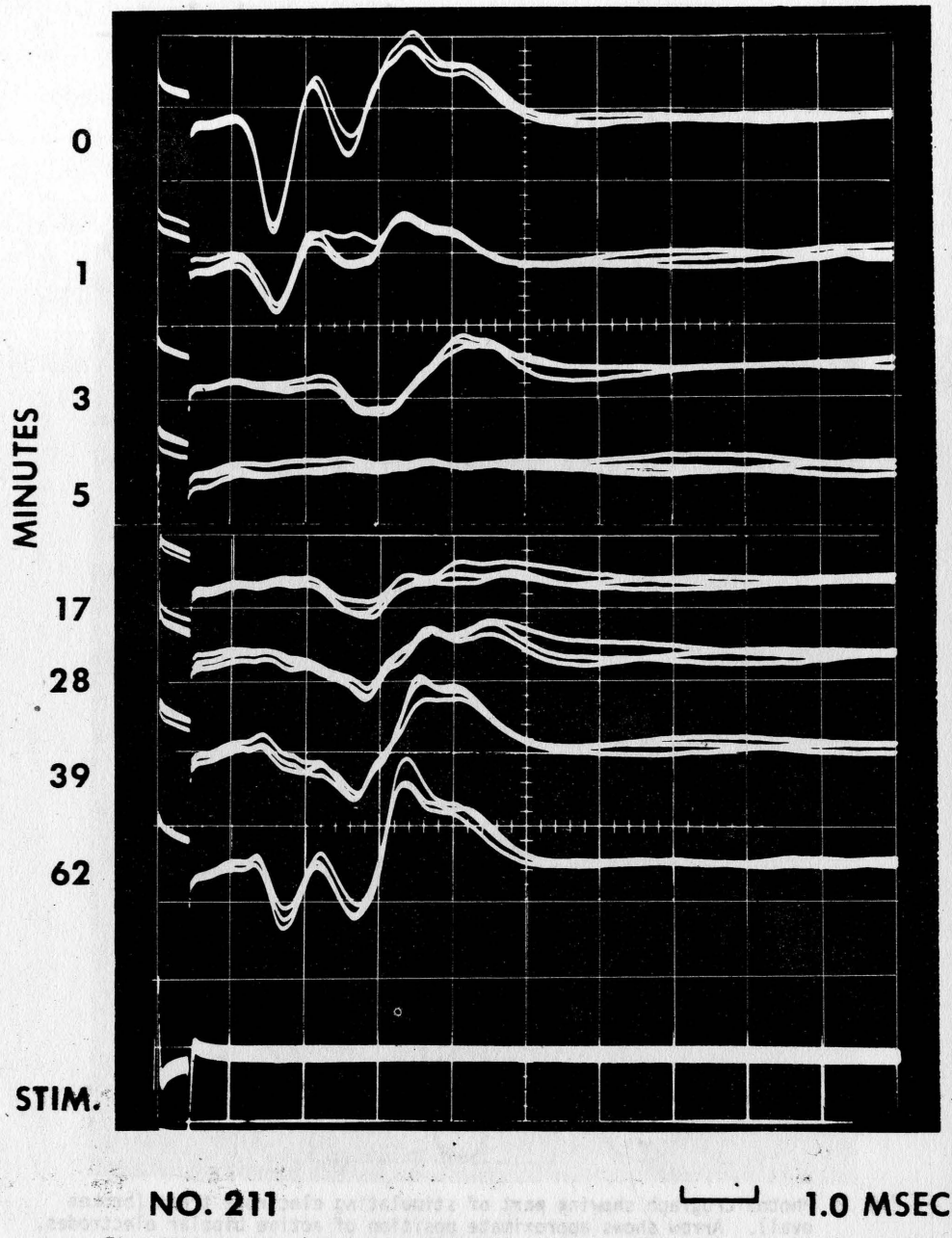


Fig. 3 Superimposed traces of medial rectus muscle potentials elicited by single pulse stimuli applied just before and at various times after the application of a neuromuscular blocking agent.

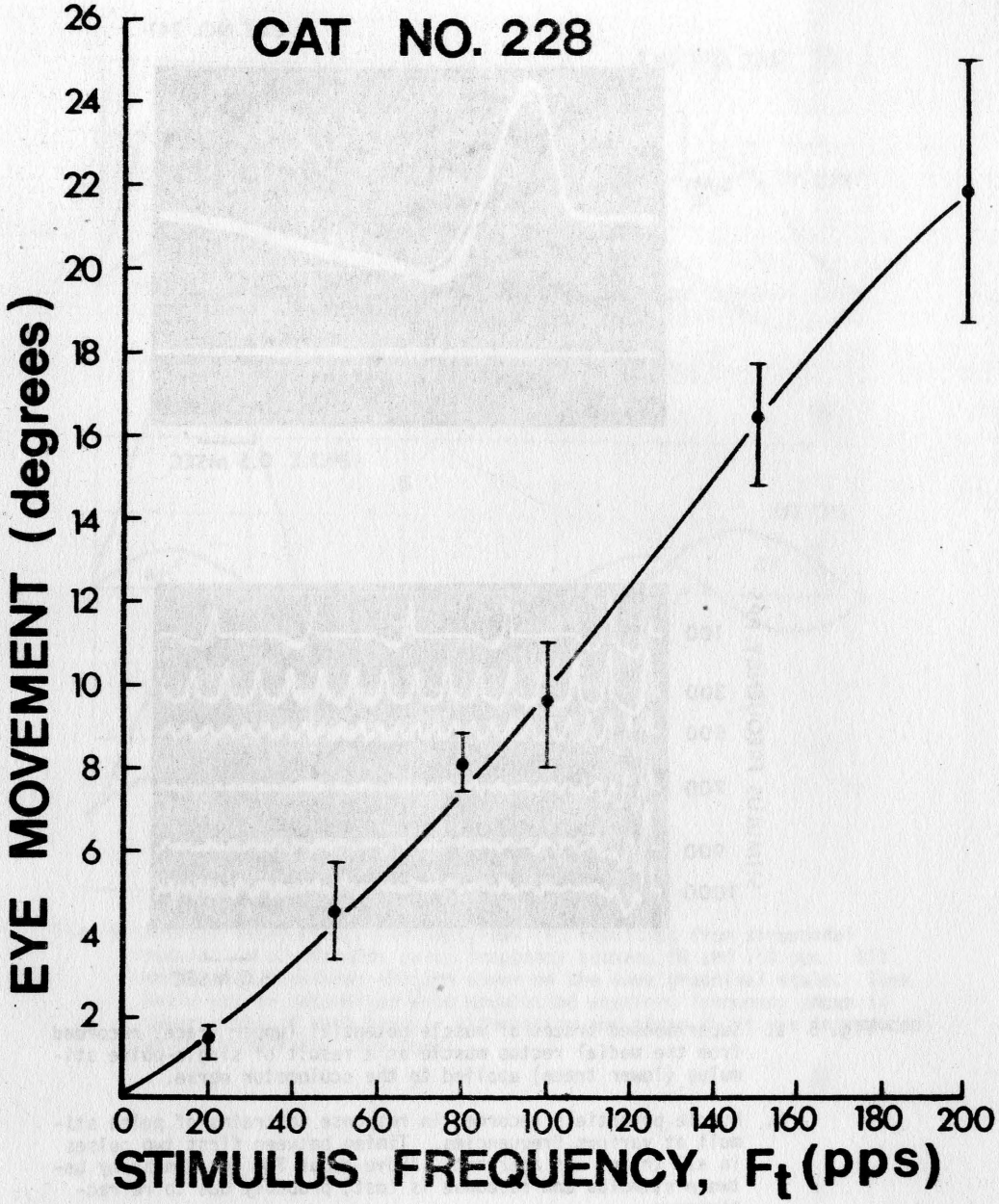


Fig. 4 Steady stat position of the eye relative to the primary position (ordinate) as a function of pulse stimulus frequency (abscissa). Vertical lines show range of data points.

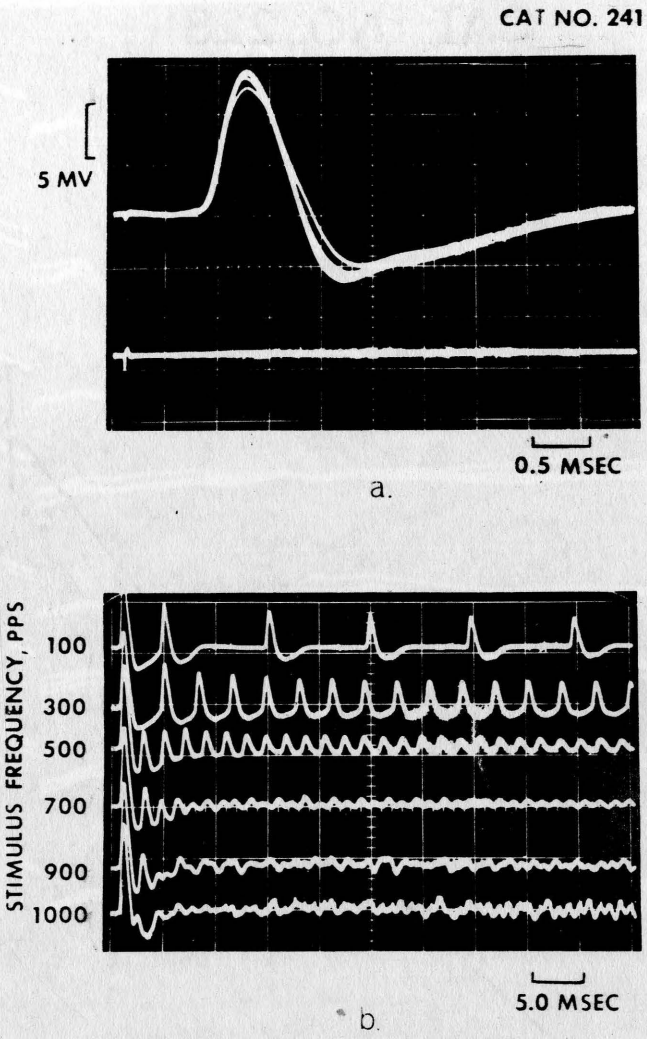


Fig. 5 a. Superimposed traces of muscle potential (upper trace) recorded from the medial rectus muscle as a result of single-pulse stimulus (lower trace) applied to the oculomotor nerve.

b. Muscle potentials recorded in response to trains of pulse stimuli at various frequencies. Timing between first two pulses in all trains was aberrant. Above about 700 pps synchrony between stimulus and response is lost, probably due to refractoriness.

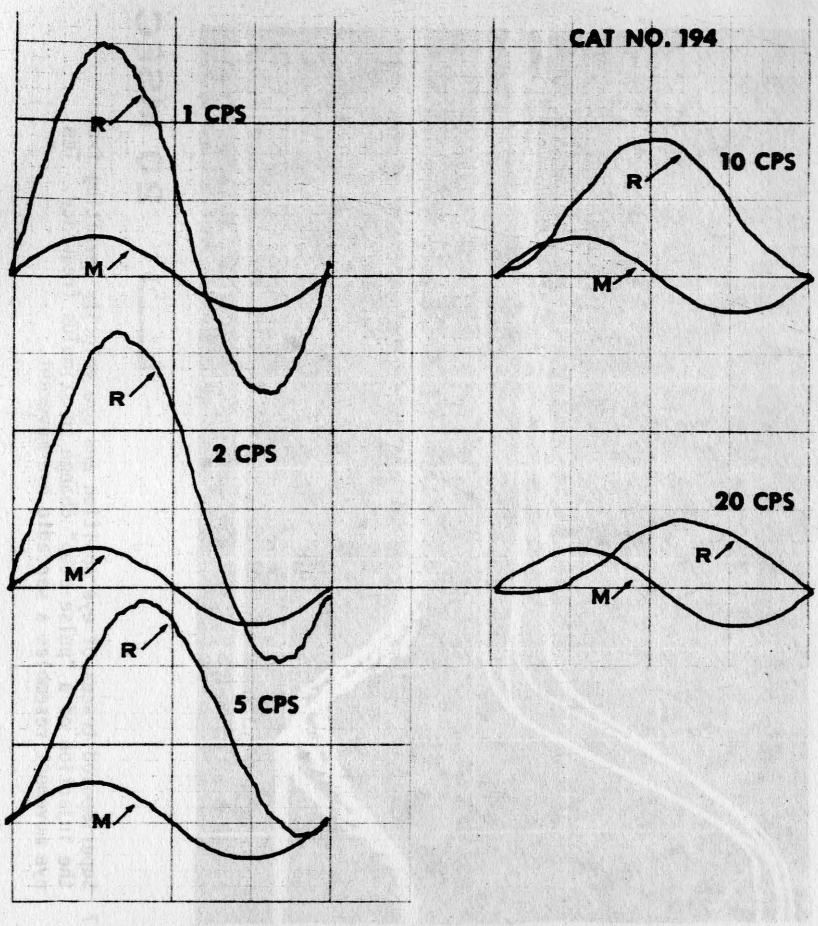


Fig. 6 Average integrated electromyograms (R) resulting from sinusoidal modulation of stimulus pulse frequency between 50 and 150 pps. All modulation waveforms (M) are shown on the same graphical scale. Time scale may be determined from modulation waveform frequency shown in upper right of each trace. Attenuation and phase shift are introduced by the integrator (see text).

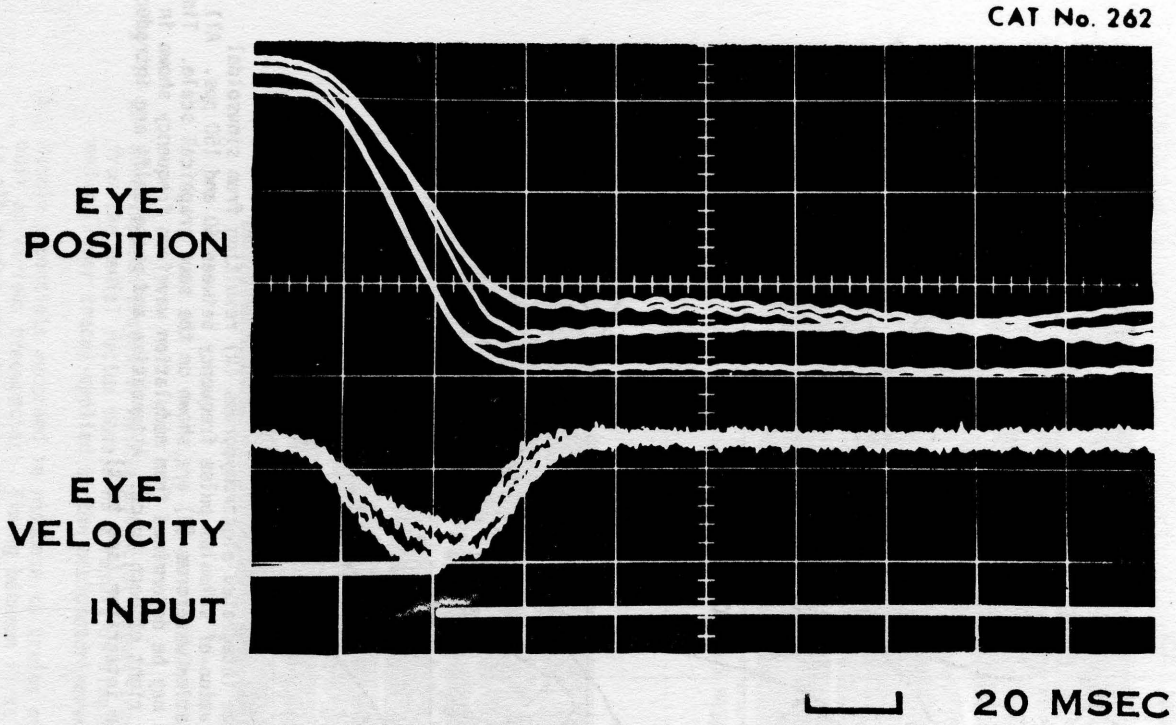


Fig. 7 Superimposed traces of eye position and eye velocity resulting from the injection of a "pulse-step" change in stimulus frequency. The eye movement resembles a saccadic eye movement.