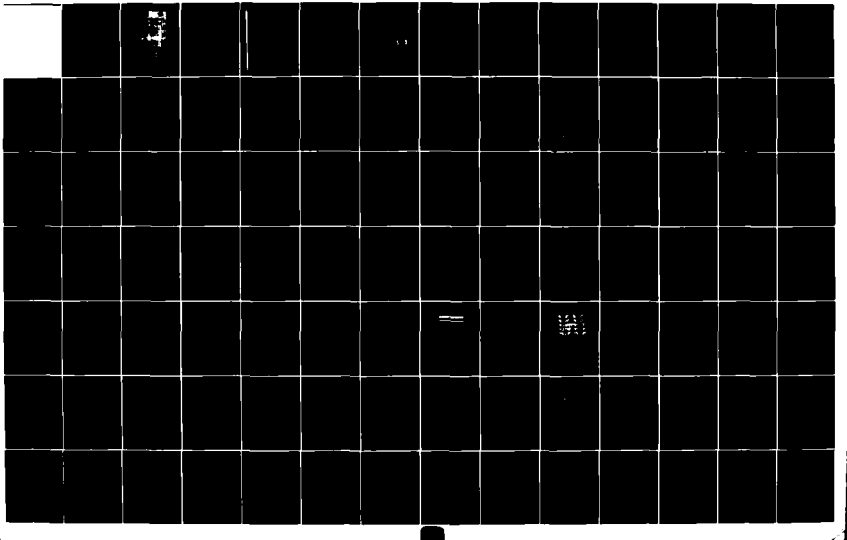
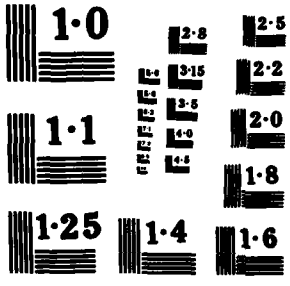
AD A145 712 MATHEMATICAL MODELS OF THE CIRCADIAN SLEEP-WAKE CYCLE (U) HARVARD MEDICAL SCHOOL BOSTON MA DEPT OF PHYSIOLOGY AND BIOPHYSICS M C MOORE-EDE 01 MAY 84 1/3

UNCLASSIFIED AFOSR-TR-84-0781 AFOSR-81-0133 F/G 6/16 NL




1·0
1·1
1·25
1·4
2·8
3·15
3·5
2·5
2·2
2·0
1·8
1·6

Mathematical Models of the Circadian Sleep-Wake Cycle

ED BY
tin C. Moore-Ede
es A. Czeisler

Mathematical Models of the Circadian Sleep-Wake Cycle

Editors

Martin C. Moore-Ede

Department of Physiology and Biophysics
Harvard Medical School
Boston, Massachusetts

Charles A. Czeisler

Neuroendocrinology Laboratory
Division of Endocrinology
Department of Medicine
Harvard Medical School
Brigham and Women's Hospital, and
Division of Health Policy Research and Education
Harvard University
Boston, Massachusetts

Circadian clocks and the control of the sleep-wake cycle is a rapidly evolving area of the biomedical sciences. There is a growing realization of the enormous importance of neural circadian pacemakers in such areas as clinical medicine, occupational health (e.g., shift-work scheduling), and transmeridian travel (i.e., jet lag). Comprehensive in scope, *Mathematical Models of the Circadian Sleep-Wake Cycle* makes available a variety of modeling tools for researchers and investigators concerned with the regulation of the sleep-wake cycle.

Recent research has identified endogenous oscillating systems in the body that control circadian sleep-wake, temperature, and hormone patterns. This finding has led to a search for models and analogies to describe the mechanisms underlying circadian rhythms. Circadian physiologists have been turning to oscillator theory, long the province of engineers and mathematicians, to aid in the conceptualiza-

(continued on back flap)

Raven Press
1140 Avenue of the Americas
New York, New York 10036

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 84-0781 2. GOVT ACCESSION NO AD-A145 712	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) MATHEMATICAL MODELS OF THE CIRCADIAN SLEEP-WAKE CYCLE	5. TYPE OF REPORT & PERIOD COVERED Final Scientific Report 5/1/81 - 11/30/82
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Martin C. Moore-Ede, M.D., Ph.D.	8. CONTRACT OR GRANT NUMBER(s) AFOSR 81-0133
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Physiology and Biophysics Harvard Medical School 25 Shattuck Street, Boston MA 02115	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2312/A1 61102F
11. CONTROLLING OFFICE NAME AND ADDRESS AIR FORCE OFFICE OF SCIENTIFIC RESEARCH /NL Bolling Air Force Base, Washington DC 20332	12. REPORT DATE May 1, 1984
	13. NUMBER OF PAGES 5
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Circadian Rhythms
Phase-Resetting of Circadian Pacemakers
Mathematical Modeling
Jet-Lag
Sleep-Wake Cycles

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This contract funded a Satellite Symposium on the Mathematical Modeling of Circadian Systems which was held on June 21, 1981 in conjunction with the Annual Meeting of the Association for the Psychophysiological Study of Sleep (APSS) from June 17-21, 1981, at Dunfey's Hyannis Hotel on Cape Cod, Massachusetts. The Satellite Symposium brought together the leading investigators concerned with modeling the circadian system to ensure that the various proposed models were critically reviewed and their strengths and weaknesses in predicting (Ctd.)

DD FORM 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

over

periodic biological phenomena were fully understood. The papers of each participant and an edited transcription of the discussion were published as a book entitled "Mathematical Models of the Circadian Sleep-Wake Cycle" by Raven Press in 1984. The published volume serves as an important source for all those who are concerned about the temporal organization of human and animal behavior and physiology.

Mathematical Models of the Circadian Sleep–Wake Cycle

Editors

Martin C. Moore-Ede, M.D., Ph.D.

Department of Physiology and Biophysics
Harvard Medical School
Boston, Massachusetts

Charles A. Czeisler, Ph.D., M.D.

Neuroendocrinology Laboratory
Division of Endocrinology
Department of Medicine
Harvard Medical School
Brigham and Women's Hospital
Boston, Massachusetts
and
Division of Health Policy Research and Education
Harvard University
Boston, Massachusetts

Raven Press ▪ New York

Raven Press, 1140 Avenue of the Americas, New York, New York 10036

Made in the United States of America

Library of Congress Cataloging in Publication Data
Main entry under title:

Mathematical models of the circadian sleep-wake cycle.

Includes bibliographical references and index.
1. Sleep. 2. Wakefulness. 3. Circadian rhythms—Mathematical models. I. Moore-Ede, Martin C. II. Czeisler, Charles A. III. Title: Sleep-wake cycle.
QP425.M35 1984 612'.022 83-19054
ISBN 0-89004-843-6

The symposium and volume preparation were in part supported by NIH Grant RR-01329 and AFOSR Grant 81-0133.

Preface

It may seem presumptuous to attempt to model so fundamental a human behavior as sleep. The time when one falls asleep appears to be so complexly dependent on subjective decisions (how absorbing a book one is reading), on prior behavior (how hard one worked that day), and on one's constitutional predilection as a "morning" or "evening" person that the very idea of mathematical description may seem preposterous to all but the most foolhardy.

Yet, what leads us to resort to mathematical modeling are the data themselves: the striking regularity of the circadian rhythms in sleep and wakefulness, body temperature, hormone levels, and many other functions especially in environments where human subjects have no knowledge of the time of day. Even more intriguing are the characteristic but complex patterns in the timing of sleep episodes that can develop in environments devoid of 24-hr time cues. Different physiological variables may display different, "free-running," non-24-hr periods that interact within the same individual, providing glimpses of internal counterpoint within the human body.

In 1972, a cluster of neurons in the hypothalamus, the suprachiasmatic nuclei (SCN), was identified as a key pacemaker of the mammalian circadian timing system. When the SCN are destroyed, circadian rhythms in a variety of physiological and behavioral functions are lost. This finding has stimulated a rapid increase in research activity on the anatomy and physiology of the circadian timing system. It has become apparent that it is a multioscillator system, with oscillators in different tissues coupled by neural and endocrine pathways.

The identification of endogenous oscillating systems within the body has led to a search for useful analogies to aid in the conceptualization of possible mechanisms that could account for the biological phenomena being observed. Hence, circadian physiologists have become increasingly interested in oscillator theory—a subject that has long been the province of the mathematician and engineer. Just as the engineer has a need to understand oscillations in complex systems, so too does the biologist; yet, there is very little in the biological literature to aid in this effort.

Attempts to model circadian systems have been made since 1960. However, only a few individuals had made serious theoretical efforts until 1973, when Pavlidis wrote *Biological Oscillators: Their Mathematical Analysis.* Since that time, there has been an intensification of interest in oscillator models that can describe circadian phenomena. Recently, books by Winfree, Enright, and Wever and articles by each of the other contributors to this volume have presented coupled-oscillator models of the circadian timing system.

These works have represented major syntheses of oscillator theory and physiological evidence, with attempts to develop mathematical models of the circadian timing system that will help define experimental questions and conceptualize the potential mechanisms that may account for the behaviors being observed. It is now possible to model many aspects of periodic human and animal behavior. Indeed, any model of a physiological system that does not take into account the system's periodic nature may have major limitations.

The author of each of the mathematical models of the circadian timing system that has been proposed has typically presented a prima facie case with little consideration of other modeling attempts. When each model has been presented at a scientific conference, it usually has been presented to an audience that has not included other investigators who have modeled circadian systems. Although each model has attracted much interest from biologists who are concerned with the strengths and the failures of prediction of the models, usually there never has been more than one proponent of a circadian mathematical model at any given meeting. Thus, it has been impossible to get a productive interaction and meaningful debate at such meetings, particularly because the mathematical subtleties are not readily appreciated in a brief presentation.

Hence, the modeling of circadian systems has been an isolated activity, with none of the normal interaction that should occur between those who have thought most about a scientific problem. The reason for this is that the proponents of the models belong to different scientific disciplines and normally never meet at national or international meetings. The Satellite Symposium at the Association for the Psychophysiological Study of Sleep, which forms the basis for this volume, brought together the various investigators who have developed models of the circadian system and allowed them to interact in a productive environment where there were also many circadian biologists who could help focus the discussion as to whether or not the various models accurately depicted the research data gathered in actual experiments.

This volume provides a review of the state of the art of circadian modeling. Discussions at the end of every chapter also provide critical insights into the strengths and weaknesses of each approach. The reverberations of the debate will be heard for many years, and this book should provide a stimulating starting point for all those who wonder what determines when we sleep and when we wake.

Martin C. Moore-Ede
Charles A. Czeisler
Boston, Massachusetts

Contents

Contributors

Domien Beersma
Department of Biological Psychiatry, Oostersingel 59, 9713 EZ Groningen, The Netherlands

Serge Daan
Zoologisch Laboratory der Rijksuniversiteit te Groningen, Kerklaan 30, 9751 NN Haren (Gr.), The Netherlands

Gerhard Dirlich
Max-Planck-Institut fuer Psychiatrie, Kraepelinstrasse 2-10, 8000 Munich 40, Federal Republic of Germany

Charmane Eastman
University of Chicago Sleep Laboratory, 5741 South Drexel, Chicago, Illinois 60637

J. T. Enright
Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093

Richard E. Kronauer
Division of Applied Sciences, Harvard University, Pierce 324, Cambridge, Massachusetts 02138

Rütger A. Wever
Max-Planck-Institut fuer Psychiatrie Klinik, Arbeitsgruppe Chronobiologie, Aubenstelle Andechs D 8138, Erling-Andechs, Federal Republic of Germany

Arthur T. Winfree
Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47906

Mathematical Models of the Circadian Sleep-Wake Cycle, edited by M. C. Moore-Ede and C. A. Czeisler. Raven Press, New York © 1984.

Mutual Excitation of Damped Oscillators and Self-Sustainment of Circadian Rhythms

J. T. Enright

Scripps Institution of Oceanography, University of California, La Jolla, California 92093

When an animal is placed under constant environmental conditions that are appropriate for the expression of endogenous circadian rhythmicity, the rhythm will usually persist indefinitely without any appreciable damping. This property of circadian systems, known as self-sustainment, is so general that when a circadian rhythm is seen to damp out under a given set of circumstances, one is apt to suspect that the wrong experimental conditions have been chosen, i.e., too cold, too much light, inappropriate monitoring equipment, and so on. It therefore seems entirely appropriate that the various sorts of single-oscillator models that have been proposed for circadian systems have embodied the capacity for self-sustainment as an essential characteristic. From a mathematical point of view, this represents a strong assumption; it focuses our attention on a relatively restricted set of oscillatory systems to the neglect of the much broader class of mechanisms and processes that can give rise only to damped oscillations.

As the modeling of circadian systems has become more sophisticated, many researchers have been led by their experimental data to propose two-oscillator models. In this case as well, it seems perfectly natural to assume that at least one of those oscillators is capable of persistent, self-sustained rhythmicity (2,10,11). Building on that tradition, it has also become customary, when proposing formulations for larger ensembles of mutually coupled oscillators, to assume that each element in the array has the capacity for self-sustainment (4,5,8,14,15). The question this article addresses is whether or not that assumption remains necessary in a multioscillator model. Suppose, instead, that each oscillator of a mutually coupled group, if it could be observed in isolation, would show only strongly damped rhythmicity: Given an impulse that sets it in motion, its rhythm will completely decay within a few cycles. Suppose, further, however, that when an oscillator is in resonance with the mutually synchronized activity of other elements in the ensemble, it receives a "push" that enhances its amplitude. Can the ensemble then show self-sustained rhythmicity? This question has both relevance and importance for circadian systems, but I will not

initially invoke any evidence whatever on that point, postponing such matters to the Discussion.

METHODS: COUPLED STOCHASTIC SYSTEMS

In order to address the question of interest, I begin with a class of coupled-oscillator models that has proved useful in other contexts and propose additional assumptions in order to speculate, "What if things were built one way rather than another?" The initial objective of these models, which I have called coupled stochastic systems, was to determine whether or not mutual coupling among oscillators can provide a plausible explanation for the precision often observed in circadian systems. Because the models and simulations have been described in detail in a recent monograph (4), only a sketchy and qualitative explanation is necessary here.

Consider first a group of mutually triggered relaxation oscillators, for example, an array of elements, each consisting of a capacitor, a voltage source, and a neon glow tube, with these units so interconnected that as soon as any one element discharges, the whole ensemble is triggered and reset to phase zero. If stochastic variability is small, the whole array will flash along at the frequency (and the level of regularity) dictated by the fastest element present. If stochastic variation is large, the role of leadership will be exchanged on a cycle-to-cycle basis among the faster elements, and such a system can behave somewhat more regularly than any single element; however, as shown elsewhere (5), this is not a particularly efficient way to gain temporal reliability from sloppy components, unless one is willing to provide an inordinate number of elements.

Because precision of system output was of central interest, I have elaborated on this scheme. Consider now an array of relaxation-oscillator-type elements in which group triggering arises not because of the first element to discharge but because of the nth element, where n is some appreciable fraction of the entire ensemble. It is easy to envision interconnections by which first-element triggering could be achieved, but how might nth-element triggering be accomplished? An engineer might well propose a scheme involving counting, but for a physiologist a more plausible alternative is to assume that the individual elements do not complete their discharges instantaneously, but instead continue to discharge over some appreciable fraction of each cycle. Then the sum of the outputs, from all active elements, can serve as the stimulus that, when of sufficient magnitude, triggers the other nondischarging elements.

This is, in very schematic outline, the kind of model, the sort of mutual coupling in which I have been interested. One of the important merits of such models is that they represent an efficient way of improving precision of the system, even when based on very sloppy oscillators (5). In order to gain some qualitative insight into how such a system functions, let us briefly look into the question of which oscillators in the ensemble will be entrained, given that they

have a broad range of intrinsic periods. It turns out that only the lower-frequency elements oscillate with a single, common average period; only the slower units are fully entrained by mutual interactions. Let us suppose that there are N elements in the array; define threshold for the triggering interaction as n_θ, a number that is an appreciable fraction of N; and set all elements into motion at the start of their charging phase. The higher-frequency elements will eventually start to discharge spontaneously, but from n_1 up to n_θ there is no interaction; only those from n_θ to N will be accelerated in that cycle. In their next cycle, the high-frequency elements will discharge even earlier, relative to the main group, and the long-term result is that the high-frequency elements "scan" repetitively through the oscillations of the mutually entrained ensemble rhythm.

Two other aspects of my simulations also deserve clarification: (a) A large measure of stochastic variability has been incorporated into the cycle-to-cycle behavior of each element. Any single element has an intrinsic period, $\bar{\tau}_i$, that defines its cycle length averaged over many cycles, but even in isolation from the ensemble, a very sloppy performance is assumed. Randomly timed events can greatly alter the realized cycle length of an element. (b) Instead of assuming that triggering of all nondischarging elements is an immediate consequence of discharge by a suprathreshold number of elements, I have treated the interaction in a probabilistic manner, which takes into account also the phase of the responding element, i.e., the time since its last discharge (Fig. 1). Extensive

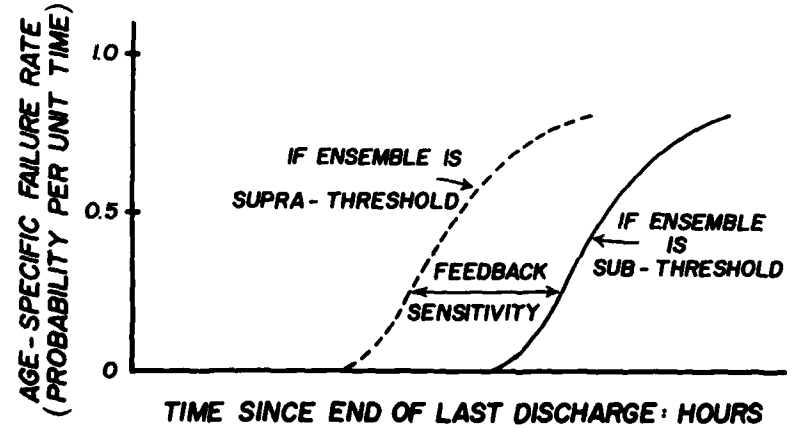


FIG. 1. Probabilistic treatment of ensemble interactions. Age-specific failure rate (a term from renewal theory) is the probability per unit time that an element will begin to discharge, given that it has not yet begun; it is here plotted as a funtion of the time since the end of its last preceding discharge. Either of the two functions illustrated will lead to a Gaussian distribution of intervals, with mean interval determined by position along the abscissa. Greater stochastic variation in performance would be associated with curves with lesser slopes. Feedback sensitivity corresponds to the parameter ϵ in the simulations leading to Fig. 2.

background from neurophysiology, involving concepts like absolute refractory time and relative refractory time, underlies this choice of formulation (4), but for present purposes, we need only recognize that this is the calculational scheme invoked by the models.

Although a verbal description of such a coupled system embodies simple-sounding ideas, translation of these ideas into a quantitative model for computer simulation requires a surprisingly long list of parameters—seven in all. Even to specify the simple concept of "an ensemble of sloppy circadian oscillators" requires four parameters: (a) the number of elements in the ensemble, (b) their overall mean period length (i.e., a scaling factor), (c) a coefficient of intrinsic variability to quantify differences among elements in their average periods, and (d) a coefficient of stochastic variability to quantify the intraelement, cycle-to-cycle unreliability. Three more parameters are required for the sort of *n*th-element coupling envisioned here: a charge–discharge factor that determines the fraction of its total cycle length that an element will be discharging; a threshold value, the equivalent of n_θ described earlier, at which interaction arises; and the magnitude of this interaction, designated "feedback sensitivity" in Fig. 1. (In most of my simulations, I have, for the sake of realism, incorporated an additional parameter, so that the elements will differ from each other not only in mean period but also in the fraction of the cycle devoted to discharge; however, that parameter is largely a luxury. I have been unable to discern significant ways in which it qualitatively alters system performance and have done many simulations without it.)

These seven parameters represent the skeleton for a coupled stochastic system. With the supplementary assumption that light alters threshold (n_θ), the models prove to be extremely versatile as descriptors for many sorts of circadian rhythm data involving responses of higher vertebrates to light regimes, as summarized elsewhere (4). This is the kind of model I have used to examine the question as to whether or not the constituent elements of a mutually coupled ensemble must be assumed to be self-sustained oscillators. To rephrase the question initially posed, "Can self-sustained rhythms in system performance arise even if each of the constituent elements is, by itself, only capable of damped oscillations?"

RESULTS

Simulations have demonstrated that the answer to the foregoing question is "yes." The initial incorporation of damped-oscillator behavior into the models involved a very modest change in formulation. The generalized models assume that each element, when discharging, contributes a value of 1.0 to a sum, and that interaction arises when this sum is greater than threshold. Assume now, instead, that each element, if examined in isolation, will have an output during its discharge phase (recognized as its potential contribution to the sum) of 1.0 in the initial cycle in which it receives a triggering impulse, a value of k in the next

cycle (where $k < 1$), a value of k^2 in the next, and so on—a process equivalent to an exponential decay of amplitude. Further, assume that if an element has been triggered by ensemble activity (i.e., has had its phase reset) during the time when most resetting occurs (within, say, the first hour after the system reaches threshold), then that element receives a stimulus that forces its amplitude back to the initial level, so that its discharge can again contribute 1.0 to the sum. The consequence of this treatment is that when any element is brought to discharge in phase synchrony with most other elements in the group, and is therefore in full resonance with the ensemble, it is not only reset in phase but also reset in amplitude, to a maximum value of 1.0.

Incorporation of this rule into simulations with coupled stochastic systems results in ensemble rhythms that show some initial damping; however, if k (the cycle-to-cycle damping factor) is not too large, the ensemble soon reaches a steady state in which its rhythm persists indefinitely with constant amplitude. Examples of the resulting trends in "amplitude" of the system oscillation are shown in Fig. 2. From these simulation data, it appears that cycle-to-cycle damping of 50% is just barely tolerable for self-sustainment of the system rhythm, but this critical value depends on several other parameters of the model. Somewhat greater damping (smaller values of k) would be acceptable if any of four parameters were to be decreased in value: stochastic, intraelement variability in period; intrinsic, interelement variability in average period; threshold for onset of feedback; or duration of the phase at which amplitude resetting occurs. The important point of Fig. 2 is only that the intrinsic damping of the constituent oscillators can be surprisingly large, provided that resonance with the ensemble provides a strong impulse and thereby restores large amplitude to the elements.

There are many obvious ways in which the simulations of Fig. 2 are unrealistic in the extreme. For example, a damped oscillator that is *almost* in resonance with the system should perhaps receive a large measure of amplitude excitement; I have given it none unless it is in nearly perfect resonance. Also, a damped oscillator that is in full antiphase with the ensemble should perhaps be damped to an unusual extent in that cycle; I have not incorporated that feature into the simulations. It is my opinion that these particular refinements would not alter the qualitative conclusions, but this interpretation rests on intuition rather than empirical demonstration.

One other matter of realism has, however, been of concern (arising from discussions with Dr. A. Winfree). For the simulations of Fig. 2, the phase-shifting effect of the ensemble on each element ("feedback sensitivity" in Fig. 1) was kept constant, regardless of that element's concurrent amplitude. This has the implausible consequence that even if an element were, through successive cycles, to reach a point of vanishingly low amplitude, it would be no more susceptible to phase resetting than when at full amplitude. I have therefore examined the consequences of several alternative formulations that seem somewhat more realistic in this regard and incorporated the idea that sensitivity

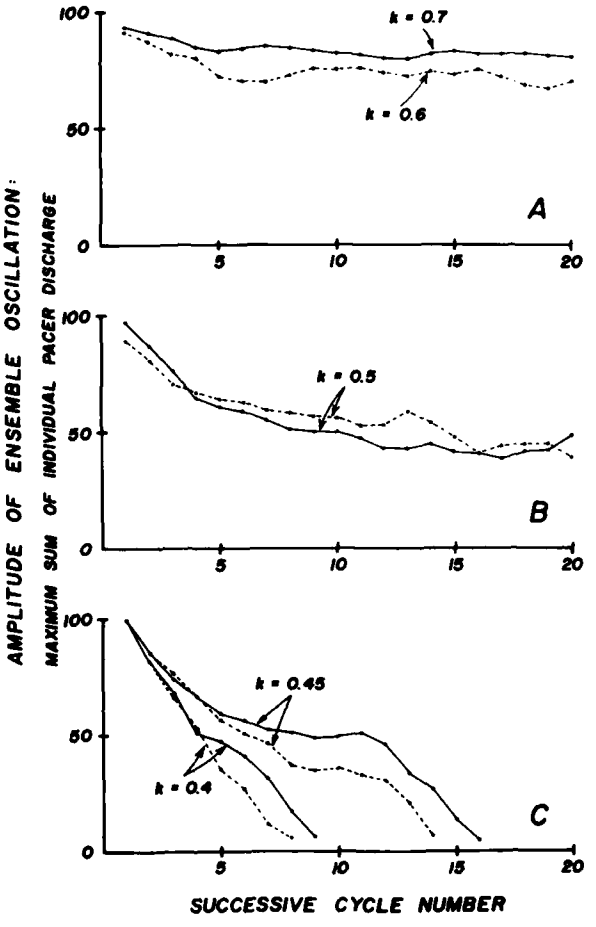


FIG. 2. Cycle-specific amplitude of ensemble oscillation (range between maximum and subsequent minimum of summed system discharge) from simulations of a coupled stochastic system in which discharge output of an element was assumed to damp in successive cycles unless discharge began in complete resonance with the ensemble rhythm. Parameter values, as defined elsewhere (4), were assigned as follows: ATU = ¼ hr; $N = 100$; $\bar{X} = 17$ hr; $\alpha = \beta = 1$ hr; $\delta = 0.5$; $\gamma = 1/24$; $\varepsilon = 8$ hr; $\theta = 0.2N$; values for k (the per cycle damping factor for output of each element) as indicated. Amplitude was reset to 1.0 if the element began to discharge during the first hour after the system reached threshold. In parts B and C, the pairs of solid and broken lines with the same k values refer to separate simulations with identical parameters, with differences in outcome resulting from stochastic factors.

of an element to phase resetting by the ensemble is inversely related to that element's concurrent amplitude. The left side of Fig. 3 illustrates the formulations investigated, and the right side shows the consequences of these formulations on trends in amplitude of the system oscillations. The data of Fig.

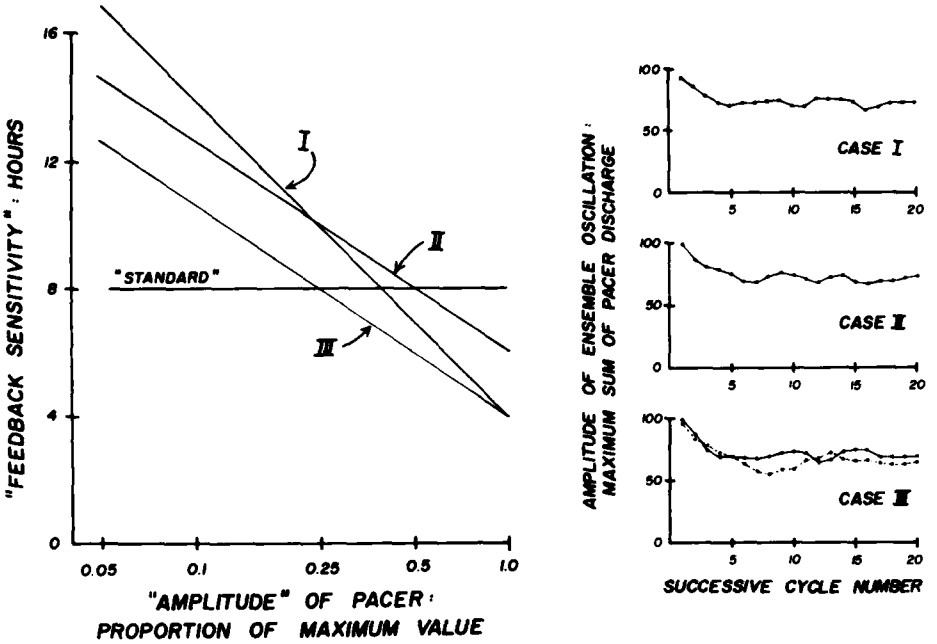


FIG. 3. Left: Three alternative formulations of an inverse relationship between feedback sensitivity, as defined in Fig. 1, and amplitude of an element, measured as its contribution to the sum of system discharge. **Right:** Cycle-specific amplitude of ensemble oscillation resulting from these formulations. All parameters except feedback sensitivity (ε) were assigned the same values as for the simulations of Fig. 2, with $k = 0.5$. Two independent simulations of case III are shown by solid and broken lines.

3 were obtained with the damping factor k set at 0.5, and the resulting rhythms were even more robust than under the initial assumptions, in the sense that there was less damping of the system oscillation (cf. Fig. 2, for $k = 0.5$).

DISCUSSION

Theory

My primary purpose in the project described here has been to focus attention on an assumption that underlies most current thinking about multioscillator models for circadian systems and to demonstrate by means of some examples that this assumption may be unnecessary. Demonstration by counterexample is, of course, an approach with many pitfalls. It is conceivable that the undamped rhythms of Figs. 2 and 3 depend critically on some unrealistic aspect of the way in which I have formulated the simulations, or on some peculiar property of

coupled stochastic systems—but I think not. Basically, the results do not now seem to be at all counterintuitive, and they have, in fact, been qualitatively anticipated by Pavlidis (9, p. 326). Suppose that the individual oscillators of an ensemble, when left to themselves, had rhythms that were subject to very weak damping tendencies (say, with $k = 0.98$), and suppose further that when coupled together, this damping process could be overcome by some very generalized sort of intense resonance effect. In this situation, regardless of the details by which the coupling mechanism produces mutual entrainment of the ensemble, it should not be surprising to find that the system could show undamped rhythmicity, provided that the restoring force at resonance is sufficiently large. Here I have represented that restoring force as immediate saturation, given complete phase synchrony; the only somewhat unexpected aspect of the results is that the potential damping of the individual elements can be quite large and still permit undamped rhythmicity of the ensemble.

Applications

Why should the issue of damped versus undamped oscillators be of interest to those studying circadian rhythms? As emphasized elsewhere (3), one of the properties of circadian rhythms that is most puzzling for an evolutionist is their self-sustainment. Dozens of ways can be easily imagined in which a daily timing ability might be useful to an organism, and there are, in addition, a variety of situations imaginable in which a *rhythmic* timing ability that persists for three or four cycles might be of ecological value. But the evolutionist has no answer when asked about the adaptiveness of a rhythm that persists indefinitely under constant conditions. As I admitted more than 10 years ago, " . . . to propose in concrete terms some plausible selective advantage which could account for this persistence is an unmet challenge" (3, p. 236).

The reality of the phenomenon is unquestionable, but the full ecological needs for endogenous timing of behavior and physiology appear, in principle, to be potentially soluble by a damped circadian rhythm, which could be regularly entrained, and even reinitiated, by the daily environmental cycle. Because an ecological need for self-sustainment is so difficult to imagine, one would seem forced to the interpretation that this property is only an accidental by-product of the way in which the rhythms are generated. Apparently, one must assume that those physiological mechanisms that can result in a circadian rhythm that functions reliably for three or four cycles will automatically lead to self-sustained rhythms (at least under some sorts of constant conditions), but "this makes severe demands on our credulity" (3, p. 236). A great many different biophysical and biochemical feedback systems can be proposed that would have oscillatory output; of these, only a very restricted subset would result in self-sustained rhythms. Consider, for example, a generalized second-order feedback loop shown schematically in Fig. 4, for which many physiological counterparts

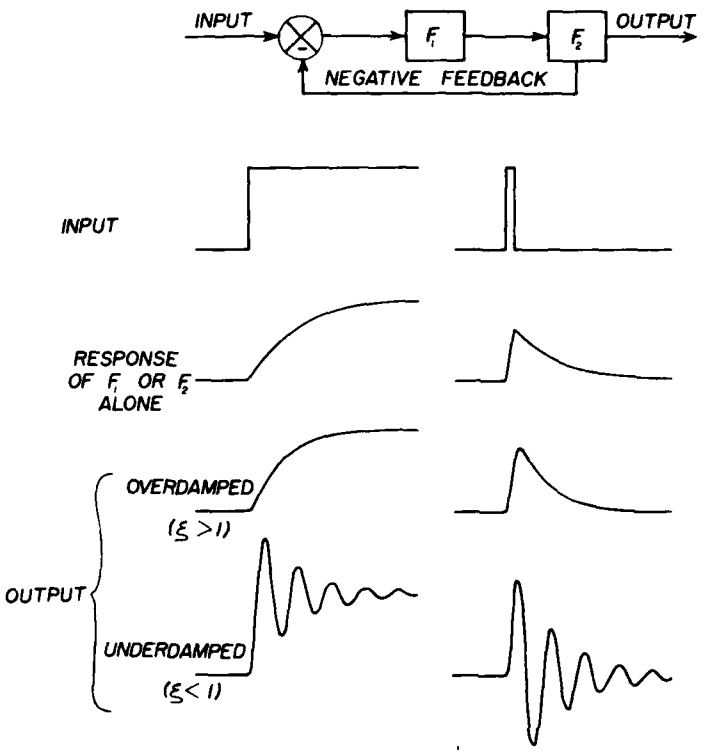


FIG. 4. Schematic diagram of a second-order linear system that can produce damped oscillatory output. Two low-pass filters, F_1 and F_2, are connected in series. Sinusoidal oscillations result from step or pulse input, provided that the damping coefficient (ξ) is less than 1.0, meaning that $A_1A_2 < (a_1 - a_2)^2/4$, where a_1 and a_2 are the rate constants of the two filters and A_1 and A_2 are proportional to their gains. Magnitude of output to the brief pulse is plotted with fourfold vertical exaggeration in the "F_1 or F_2" case and the overdamped case.

can easily be imagined, e.g., sequential, self-inhibiting chemical reactions. Neither of the low-pass filters alone will lead to oscillations, but if the two are connected in series in such a loop, the system will produce damped sinusoidal oscillations with constant period following pulsed input, over a broad range of parameter values for the filters. Even the sequence of the two filters (i.e., whether the slower or the faster is first in series) is irrelevant. Note that this type of generalized system seems to be automatically excluded from consideration as a basis for circadian rhythms because its oscillations are not self-sustained. This kind of system *could*, however, provide the building blocks, the damped-oscillator type of elements, that might participate in a coupling scheme resembling the one considered here (as I have demonstrated by further simulations). The fixation on the undamped rhythm of the whole animal has automatically, and mistakenly, led to the acceptance of the idea that the

constituent oscillators of a circadian ensemble must also have the intrinsic capacity for self-sustainment.

Most researchers in the field of circadian rhythms are now willing to entertain the idea that in higher animals the pacemaker may well consist of a system of mutually entrained oscillators. Stripped to its essentials, mutual entrainment implies a sort of interaction that produces appropriate phase shifts of the individual component rhythms, so that they are kept in synchrony. All previous considerations of mutually entrained oscillators as models for circadian systems have emphasized this phase shifting, with little attention, if any, to the possibility of effects on amplitude. The result of that approach is that the puzzling phenomenon of self-sustainment remains unexplained; it is a property that then must be taken as an assumed property of the constituent oscillators. However, if one entertains the hypothesis that the phase shifting of mutual entrainment may also have systematic effects on amplitude of constituent oscillators, so that those in resonance with the ensemble are subject to amplitude enhancement, then one can potentially account for self-sustainment of the whole-animal rhythm as an emergent property associated with coupling of an array of oscillators that by themselves need not be self-sustaining.

It is worth emphasis that the required hypothesis is a very modest assumption. One can, of course, imagine schemes in which phase shifting will have no effect on the amplitude of an oscillation (4), but generalized sorts of oscillatory dynamics, such as those associated with the scheme shown in Fig. 4, will ordinarily have exactly the property hypothesized here: Any input that can shift the phase of an ongoing rhythm automatically also has the capacity to alter oscillatory amplitude.

The outcome of these considerations is the following set of propositions:

1. There are clear ecological advantages associated with circadian rhythms that will persist reliably for several cycles.
2. For a multicellular organism to achieve this goal, a plausible mechanism would be environmental synchronization of an ensemble of mutually entrained oscillators.
3. Mutual entrainment can be reasonably assumed to involve amplitude enhancement for those oscillators that are most closely in resonance with the ensemble.
4. If this amplitude enhancement is sufficiently strong, the result should be self-sustainment of the ensemble rhythm (i.e., at the whole-animal level), even if the constituent oscillators do not have this property.

Another way of summarizing these propositions is to say that once circadian rhythms are viewed as the output of a coupled multioscillator system, the observation that the rhythms are self-sustained need not be regarded as particularly surprising; but the experimentalist has every right to ask how such speculation can be useful. What is the empirical evidence in favor of or against this interpretation, and what experimental tests can be proposed to distinguish among alternatives? One important piece of evidence that deserves attention here

is the demonstration that isolated single cells of the alga *Acetabularia* are capable of many cycles of apparently undamped circadian rhythmicity under constant conditions (6). Hence, multicellular interactions are not essential to self-sustained circadian rhythms. (One cannot be fully certain, on the available evidence, that the rhythms of *Acetabularia* are truly self-sustained, as are those of higher vertebrates, but they are at least not strongly damped.) *Acetabularia* is, of course, an extremely large cell—large enough that one could speculate about an ensemble of mutually entrained, intracellular oscillatory systems and still remain within the context of the viewpoint considered here. But it remains important to recognize that circadian rhythms, which are apparently undamped, can be observed even at the level of the single cell. One cannot, therefore, legitimately ignore the possibility that the self-sustainment observed in the circadian rhythms of higher animals is simply an evolutionary inheritance from unicellular ancestors that had already developed that capacity.

There are a good many cases in the literature in which clear damping of an overt circadian rhythm has been empirically observed, cases that might be taken as evidence in favor of the conceptual scheme described here, but unfortunately these examples are usually open to alternative interpretations. If one assumes that an animal's pacemaker is a multioscillator ensemble, then damping of an overt rhythm *could* represent decreases in amplitude of a non-self-sustained circadian oscillator or group of oscillators, the sort of phenomenon postulated here. It could, however, also represent the gradual loss of synchrony within an array of self-sustained oscillators because of the absence of adequate entraining stimuli, either those stimuli leading to mutual entrainment of the ensemble or those associated with external synchronization. This problem proves to be very general; in most cases, it will be very difficult to distinguish between these two alternatives. However, the problem is not completely hopeless.

One of the experimental techniques for producing complete damping of a circadian rhythm is to expose a diurnal bird to constant bright light. The usual result of this treatment is that overt rhythmicity vanishes over a few days, during which the amplitude of the circadian cycle rapidly diminishes. This phenomenon seems to be readily interpretable (4) as indicating loss of phase synchrony among constituent elements of a multioscillator ensemble in which the elements are normally mutually entrained. A subsidiary experiment, however, indicates certain difficulties with this interpretation: The subsequent transfer of such an arrhythmic bird into constant dim light or darkness usually results in immediate recovery of rhythmicity, with phasing roughly determined by the time of the bright-to-dim transition; moreover, the recovered rhythm usually begins with its full, steady-state amplitude (1, and J. T. Enright, *unpublished data*).

If one assumes that the constituent oscillators of the hypothesized ensemble continue their undamped, self-sustained rhythms under bright light, eventually achieving a random distribution of phases relative to each other but having otherwise unchanged properties, then it is surprising to find that the restored system rhythmicity immediately recovers to its full amplitude. Instead, one would expect a gradual buildup of the whole-animal rhythm as the constituent

oscillators are gradually drawn back into full synchrony with others in the group. A possible remedy for this difficulty is to postulate that the constituent oscillators of the hypothesized ensemble are capable only of damped rhythms, once mutual entrainment is lost, and that therefore all are completely resettable to a common phase at the onset of system rhythmicity. It should be noted, however, that this sort of interpretation does not necessarily demand amplitude damping. It invokes, instead, a progressive increase, during constant light, in responsiveness to resetting stimuli, an increase in feedback sensitivity (Fig. 1), such as was proposed as a likely concomitant of damped-oscillator behavior. However, such change in sensitivity could, of course, also arise in self-sustained oscillators during constant light. Damped oscillators *could* account for the experimental results, but undamped oscillators would also be adequate, granted an ad hoc supplementary assumption.

Another potentially relevant line of evidence is contained in the data illustrated in Fig. 5, from recent elegant experiments with isolated chick pineal organs (13). In the presence of an entraining light cycle, the rhythmic output continues without significant damping for at least four cycles, demonstrating the adequacy of the culture conditions. Under constant dark, however, the rhythm shows rapid decrease in amplitude: "it was . . . heavily damped compared to the rhythm present under light/dark conditions" (13). This result is initially a disappointing one for the hypothesis (7) that the pineal organ may be the "master oscillator" of the bird circadian system, because a capacity for self-sustained rhythmicity is one of the key properties by which a master oscillator would presumably be identified. However, the recognition that an ensemble of damped oscillators could, through their coupling, produce a self-sustained rhythm suggests that the search for a master oscillator, a dominant, discretely localizable driver for the whole-animal circadian system, may be a search for a

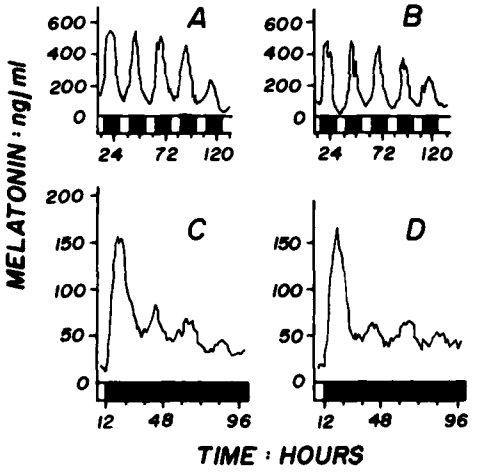


FIG. 5. Melatonin output from single, *in vitro* cultured chick pineal organs under light/dark conditions (A and B) and under constant darkness following light cycles (C and D). Note strong damping of the rhythms in darkness. The culture flow rate was twice as high for experiments of C and D as for those of A and B, which accounts for differences in amplitude in the light/dark response. (From Takahashi et al., ref. 13, with permission.)

phantom. Perhaps the interactive pacemaker *system* consists only of damped ensembles, with behavior in isolation that resembles that of the pineal organ. Hence, the bird pineal organ may turn out to be a more important component of the circadian system than the initial consideration of the results shown in Fig. 5 would lead one to suspect.

It should be noted that the results of Fig. 5 are open to at least one other alternative interpretation. The data in Fig. 5C are typical of what would ordinarily be expected of a damped oscillator, with progressive decline in amplitude over successive cycles; however, the data of Fig. 5D are somewhat different. No clear sign of further decrease in peak height was evident after the second peak in melatonin. Hence, it is not well established that the *steady-state* behavior of the organ in constant darkness should be considered a damped rhythm. Figure 5D is compatible with the idea that the amplitude of the first free-running peak is enhanced by the last-seen light treatment and that the rhythm thereafter shows no damping greater than that under light/dark conditions (K. Hoffmann, *personal communication*).

Even if one is willing to set aside such reservations and accept that the isolated pineal organ behaves as a true damped oscillator, the question remains as to whether this result indicates that cellular rhythms that are self-sustained and undamped rapidly get out of synchrony with each other in the absence of entrainment by light, as suggested by Takahashi et al. (13), or whether it indicates instead the damping of the rhythms of individual cells of the tissue. A definitive answer to this question will depend on measurements of rhythmicity in single, isolated pineal cells, a demand that goes far beyond present-day experimental techniques. Nevertheless, the implications of the desynchronization interpretation can be examined. If we assume that the rhythm of the whole organ, during entrainment and the first free-run peak, accurately reflects the fully synchronized rhythms of the single cells, an extremely broad range of free-running periods among the individual cellular oscillators would be required in order to produce damping that is as intense as that observed under constant conditions. Rough calculations based on Fig. 5C suggest that the required distribution of free-running periods would have a standard deviation on the order of 6 hr; that is, the "circadian" periods of the cells would range from about 12 to 36 hr. [Because the width of the first, presumably synchronized, melatonin peak is about 12 hr, assume that the single cell's melatonin output has a Gaussian distribution in time, with standard deviation on the order of about 3 hr. Assuming additivity of variances, the second peak would have a variance of $(9 + \sigma_\tau^2)$ hr^2, the third a variance of $(9 + 4\sigma_\tau^2)$ hr^2, the fourth a variance of $(9 + 9\sigma_\tau^2)$ hr^2, where σ_τ is the standard deviation of free-running period. Noting that the peak height of a Gaussian distribution is inversely proportional to its standard deviation, we find that if σ_τ is taken to be 6 hr, the heights of the second, third, and fourth peaks should be about 45%, 24%, and 15%, respectively, of the height of the first peak. Such values are in reasonable agreement with the data of Fig. 5C.] There are, of course, no empirical data to

indicate whether such a broad distribution of period values is either plausible or very unlikely, but it is my intuition that natural selection would have led to circadian oscillators appreciably less variable in period. Hence, an interpretation based on damped circadian oscillators is more appealing with such data as in Fig. 5 than is the idea of loss of phase synchrony among undamped oscillators.

Neither of the two examples considered here can be taken as strong evidence that the circadian systems of higher animals include strongly damped oscillators as components of a mutually coupled system, and other more persuasive evidence is not presently available. Nevertheless, I am optimistic that such evidence may well be found. Recent research has demonstrated that restricted feeding regimes, although usually unable to entrain in an animal's entire intact circadian system, are nevertheless often able to induce persistent activity patterns that are most easily interpreted in terms of entrainment of a subset of oscillatory components within the circadian system. The results of such experiments are particularly clear-cut when undertaken with animals in which the suprachiasmatic nucleus—a presumed site for the light-sensitive circadian pacemaker—has been ablated (12). However, such rhythmic patterns of activity, induced by food regimes, do not persist as self-sustained oscillations when food is thereafter offered *ad libitum*; instead, they show acute damping within a few cycles. Whether or not this damping should be interpreted as true damping of circadian oscillators or as some other phenomenon, such as loss of mutual synchrony among self-sustained oscillators, remains to be established, but this experimental system appears to be ideal for further study of this question. Perhaps the concept of damped oscillators will prove to be superfluous for the interpretation of circadian systems; nevertheless, it is a possibility that at least deserves more thorough exploration than has been attempted to date.

ACKNOWLEDGMENTS

Drs. John Thorson, A. T. Winfree, and G. D. Lange have contributed importantly to the ideas developed here. Drs. J. Aschoff and K. Hoffmann offered valuable comments on an early version of the manuscript. This research was supported, in part, by Grant PCM-7719949 from the National Science Foundation.

REFERENCES

1. Binkley, S. (1977): Constant light: Effects on the circadian locomotor rhythms in the house sparrow. *Physiol. Zool.*, 50:170–181.
2. Daan, S., and Berde, C. (1978): Two coupled oscillators: Simulations of the circadian pacemaker in mammalian activity rhythms. *J. Theor. Biol.*, 70:297–313.
3. Enright, J. T. (1970): Ecological aspects of endogenous rhythmicity. *Annu. Rev. Ecol. Syst.*, 1:221–238.

4. Enright, J. T. (1980): *The Timing of Sleep and Wakefulness: On the Substructure and Dynamics of the Circadian Pacemakers Underlying the Wake-Sleep Cycle. Studies in Brain Function*, Vol. 3. Springer-Verlag, Berlin.
5. Enright, J. T. (1980): Temporal precision in circadian systems: A reliable neuronal clock from unreliable components? *Science*, 209:1542–1545.
6. Karakashian, M. W., and Schweiger, H. G. (1976): Circadian properties of the rhythmic system in individual nucleated and enucleated cells of *Acetabularia mediterranea*. *Exp. Cell Res.*, 97:366–377.
7. Menaker, M., and Zimmerman, N. (1976): Role of the pineal in the circadian systems of birds. *Am. Zool.*, 16:45–55.
8. Pavlidis, T. (1969): Populations of interacting oscillators and circadian rhythms. *J. Theor. Biol.*, 22: 418–436.
9. Pavlidis, T. (1971): Populations of biochemical oscillators as circadian clocks. *J. Theor. Biol.*, 33:319–338.
10. Pittendrigh, C. S. (1981): Circadian organization and the photoperiodic phenomena. In: *Biological Clocks in Seasonal Reproductive Systems*, edited by B. K. Follett, pp. 1–35. John Wright, Bristol.
11. Pittendrigh, C. S., Bruce, V. G., and Kaus, P. (1958): On the significance of transients in daily rhythms. *Proc. Natl. Acad. Sci. USA*, 44:965–973.
12. Stephan, F. (1981): Limits of entrainment to periodic feeding in rats with suprachiasmatic lesions. *J. Comp. Physiol. A*, 143:401–420.
13. Takahashi, J. S., Hamm, H., and Menaker, M. (1980): Circadian rhythms of melatonin release from individual superfused chicken pineal glands *in vitro*. *Proc. Natl. Acad. Sci. USA*, 77:2319–2322.
14. Winfree, A. T. (1967): Biological rhythms and the behavior of populations of coupled oscillators. *J. Theor. Biol.*, 16:15–42.
15. Winfree, A. T. (1980): *The Geometry of Biological Time*. Springer-Verlag, Berlin.

DISCUSSION

Dr. Czeisler: Do you view the different neurons to which you referred as residing together in one location in the central nervous system, thus forming a precise pacemaker that then drives the rest of the circadian rhythms of the organism? Or do you view these separate elements as being located in different parts of the whole organism, which in its entirety produces one observable rhythmicity?

Dr. Enright: For the general purposes of the model it does not matter. My own interpretation of the data from birds suggests that the pineal organ may be a part of the pacemaker, probably a component of the discriminator which I have proposed here. But the whole pacemaker is not located in the pineal organ. I think there is strong reason to suspect the suprachiasmatic nucleus as being the site of many such oscillators which may represent the entire ensemble of the pacemakers that I have talked about here. It would not disturb me at all to entertain the idea that many of these oscillators are located in the suprachiasmatic nucleus. There may be other ensembles elsewhere. So I am simply not going to pin myself down. Sorry.

Dr. Edmunds: We were treated at a recent Gordon conference to some data showing that cockroaches entrained to particular non-24-hr periods were apparently able to free-run but retain the non-24-hr period for several weeks. Do your simulations show this?

Dr. Enright: The simulations show aftereffects comparable in direction, but not in magnitude to the phenomenon you are describing. This corresponds to the mammalian data on aftereffects. If you entrain a hamster to a 22-hr light/dark cycle, you will not subsequently see a free run that has a 22-hr period. You will see one that has perhaps a 23-hr free run, rather than the normal 24 hr. Similarly, the aftereffects that I have been able to simulate usually do not display the full "memory" of the period to which the animal has previously been subjected.

Dr. Dirlich: In your book you emphasize one crucial property of your model, that is, that the cycles are so slow they are circadian in nature, while the firing of neurons mostly is at a much higher frequency. What can you do to explain this problem?

Dr. Enright: It is indeed one of the most outrageous assumptions of the model that a single neuron can show a circadian periodicity. It is an outrageous assumption because there are no data that I know of from isolated neurons to support this interpretation. There are excellent data from unicellular organisms, algae in particular, that show that a single cell is a sufficient physiological entity for circadian rhythmicity to develop. But in terms of circadian rhythms in single cells, there are no data that are available. The original claims by Strumwasser that he had found such circadian rhythms in the single cells of *Aplysia* have, I think, largely been retracted. There is evidence for aftereffects of the prior light/dark cycle which persist for the first subsequent postoperative peak, but thereafter a persistent rhythm in a single cell has not been demonstrated.

Dr. Weitzman: Is your model really getting at the mechanism of the oscillator rather than being a model of complex multioscillator function?

Dr. Enright: That is right. It is a one-oscillator model.

Dr. Weitzman: It does not assume there is only one oscillator. You are looking at *an* oscillator and trying to understand how it works rather than showing that all behavior is based on one oscillator.

Dr. Enright: Yes. But this interpretation is justified by the animal and bird data. For nearly all experimental results, a single oscillator is sufficient to account for the data.

Dr. Weitzman: In blinded birds with feathers intact on top of the head, there was a very significant difference in the activity/rest ratio when the birds were free-running as compared to when they were entrained by a light/dark cycle. Not only did the entrainment process change, but also there was a very significant difference in the activity/rest ratio from one to the other. Now, how would a single-oscillator model explain such changes?

Dr. Enright: Without trying to go into detail, that is one of the questions that I've looked into in great detail in my simulations. I have not been concerned simply with the period, but with the distribution of activity time and rest time. I have not attempted to simulate this particular result in the blinded sparrow, but I do not think I would have any difficulty.

Dr. Kronauer: Woody Hastings has spent a lifetime studying *Gonyaulax*, a single cell with circadian rhythmicity, which at constant temperature has an accuracy in its free-running period of a few percent. So it seems to me that within a single cell you can have a very accurate pacemaker.

Dr. Enright: Hastings has, in fact, in his very beautiful work, studied entire populations of cells in a test tube, in which there remains the residual concern that there may be some interaction within the ensemble—I know that he has tried to rule that out—and so I am more impressed with the data from *Acetabularia*, where one knows one has a single cell. One finds a rhythm which persists and which shows a cycle-to-cycle variability with a standard deviation of about 2 hr.

Dr. Kronauer: Beatrice Sweeney has actually isolated single *Gonyaulax* cells.

Dr. Enright: Sweeney did a couple of experiments which were published in the Cold Spring Harbor Symposium (*Cold Spring Harbor Symp. Quant. Biol.*, 25:145, 1960). Let me simply say that I have recognized for a while that circadian rhythms can exist within a single cell. We see it unequivocally in *Acetabularia*. We see it on a very probable basis when one looks at assembled populations of unicellular organisms like *Gonyaulax* and *Euglena*.

Mathematical Models of the Circadian Sleep-Wake Cycle, edited by M. C. Moore-Ede and C. A. Czeisler. Raven Press, New York © 1984.

Toward a Mathematical Model of Circadian Rhythmicity

Rütger A. Wever

Chronobiology Laboratory, Max Planck Institute for Psychiatry, D-8138 Andechs, West Germany

The efficiency of a model describing biological phenomena, such as circadian rhythmicity, can be measured by the ratio between the basic preconditions put into the modeling process and the predictions deduced from different model applications and confirmed by the results of biological experiments. No single model can describe all the different aspects of a biological phenomenon equally well; rather, there will be several models complementing one another, each of which will describe specific aspects of the system. In order to critically examine any model, the basic preconditions used for constructing the model should be listed, and the predictions tested through biological experimentation. If possible, these predictions should be formulated quantitatively.

Models have been developed that simulate the dynamics of biological systems; these are usually "hardware" models that consist of mechanical, hydraulic, or electrical devices. Their mechanisms have, in principle, no direct relation to biological mechanisms. Such models are rarely flexible enough to simulate all biological conditions, and applications of such models run the risk of inappropriately representing biological mechanisms. Other models are based on features of the biological systems under consideration, e.g., on neuronal interconnections, properties of membranes, or structural transformations. These simplifications of the structural processes of the biological phenomena under consideration can lead to a more thorough understanding of them. However, there is a danger in such models that correlation between model predictions and experimental data may suggest a similar correlation between the structural mechanisms of the model and the biological system; knowledge of the mechanisms underlying circadian rhythmicity is as yet insufficient to apply these hardware models effectively.

Thorough understanding of the dynamics of biological processes is a precondition for subsequent analysis of their underlying mechanisms. It is therefore advantageous to use "software," or mathematical, models that describe these dynamics without assuming their anatomical basis. Of relevance here are kinetic models that render the dynamics of the system; special wave shapes are then the result of computations describing the behavior of the system

under varying external conditions. The mathematical expression of such a dynamic type of model is a differential equation, and the expressions of peculiarities in biological results are nonlinearities.

A mathematical model describing circadian rhythmicity will be presented based on a simple differential equation. The method of deducing the model is that of trial and error. Alternative models and modifications are formulated and tested, based on relevant biological experiments, in successive iterative steps (31). After establishing the relevant model equations, solutions of this model will be presented, as computed under varying external conditions, constituting predictions for the behavior of the biological system under various environmental conditions. Different types of equations have the capacity to describe correlations between separated rhythm parameters and experimental conditions when the coefficients are selected properly. More powerful is the prediction of summarizing multifold correlations and, in particular, of interdependences between different rhythm parameters in both the steady state and during transient states. Therefore, of special relevance is the determination of many different rhythm parameters in both the mathematical analysis and the biological experiment.

DEDUCTION OF THE MODEL EQUATION

The Initial Equation: Stochastic versus Oscillatory Approach

In establishing a model of circadian rhythmicity, the first step is to decide whether this model should be based on stochastic or oscillatory processes; however, the two methods lead to remarkably coinciding results. The modeling process cannot culminate in a deterministic model, because all endpoints of rhythms to be determined in biological measurements show random fluctuations; most rhythm parameters, the interdependences of which are of special concern, do not yield absolute values, but rather variabilities of rhythm endpoints.

In the analysis of long-term variations in autonomous rhythms, a negative serial correlation between the duration of successive cycles within a circadian time series is a consistent result (15,45,49). Such a correlation is achieved between a relatively stable "pacemaker" that might be of stochastic or oscillatory origin and an "overt rhythm" that is controlled by the pacemaker via relatively labile coupling processes. As an alternative, a serial correlation is obtained by a special type of oscillator that generates those correlations under the influence of random noise. Discrimination between these concepts is possible by analyzing the increase of the "relative stability," or the ratio between long-term variability (calculated as twice the standard deviation of

successive reference phases around the computed linear regression—the reciprocal value of "stability") and short-term variability (calculated as the standard deviation of the durations of successive cycles—the reciprocal value of "precision") with increasing length of the time series under consideration. Results from long-term human experiments and various animal experiments are compatible only with the oscillatory concept (56). Simultaneously, these analyses indicate that overt rhythms reflect, to a great extent, properties and variabilities of pacemakers and that the interconnecting coupling processes are rather negligible in their effects on these rhythms.

Using the oscillator instead of the stochastic process as the basis for further modeling evaluations, a differential equation is formulated, beginning with the simple oscillation equation of the second order:

$$\ddot{y} + 2\beta\dot{y} + \omega^2 y = z \tag{1}$$

In this equation, y represents the oscillating variable, which is a function of time t; $\dot{y}$ and $\ddot{y}$ are the first and second time derivatives of y. The variable z represents the "external force" controlling the oscillation, i.e., the independent variable; it may be constant or likewise may depend on time t. β and ω are free parameters. With a constant value of z, equation 1 has the solution

$$y = z/\omega^2 + A\, e^{-\beta t} \sin(\sqrt{\omega^2 - \beta^2} \cdot t + \psi) \tag{1a}$$

This solution shows that, in the long run, y approximates z/ω^2 as either oscillatory (if $0 < \beta < 1$) or aperiodically adapting (if $\beta > 1$). In case of oscillatory adaptation, the coefficient of $\dot{y}$ in equation 1, i.e., the "damping," determines the rate of fading away of the oscillation per unit time, and the coefficient of y in equation 1, i.e., the "restoring force," determines—with great values of β together with the coefficient of $\dot{y}$—the frequency of the damping oscillation; the parameters A (initial amplitude) and ψ (phase) depend on the initial conditions. In the case of $\beta = 0$, i.e., when the term with $\dot{y}$ is missed in equation 1, the oscillation remains running infinitely, with the frequency ω and a constant amplitude that is determined exclusively by the initial conditions. In the case of a negative value of β, equation 1 describes an oscillation with infinitely increasing amplitude. In the case of a negative value of ω^2, the system is not pushed back to its neutral position after every elevation but is pushed even father away, becoming unstable.

If z in equation 1 is not temporally constant but is a periodic function of time, y does not approximate a constant value, but a periodic function: z operates as a forcing oscillation that synchronizes the forced y oscillation. The frequency of the forced oscillation equals that of the forcing oscillation, independent of ω; its steady-state amplitude and the phase-angle difference between the z and y oscillations are essentially determined by the ratio between the frequency of the z oscillation and the intrinsic frequency of the y oscillation.

Nonlinear Damping

Circadian rhythms are considered to be endogenously generated, i.e., based on self-sustaining oscillations (1,13). The only alternative to this concept has been discussed by Brown (8,9). He assumed all rhythms to be products of environmental influence. Under "constant conditions" with "free-running" rhythms (having periods independent of environmental time cues), he postulated influence by "subtle stimuli" operating via frequency transformation. However, it has been shown that synchronization of a rhythm by environmental stimuli, including "subtle stimuli" (49), is effective because of phase control, not frequency control. Because phase control is not compatible with frequency transformation (46), this sole alternative to the concept of self-sustainment has been ruled out.

The next alternative concerns the initial behavior of self-sustaining oscillators. Such oscillators either can be self-excitatory, i.e., capable of starting to oscillate from rest spontaneously, or can depend on external stimuli to initiate the mechanism. In other words, the self-sustainment mechanism either is permanently in operation or is in operation only after the system is elevated above a certain threshold by an external stimulus. It is only recently that circadian rhythms have been shown to be self-excitatory (51); therefore, the modeling process must start with an oscillator that is both self-sustaining and self-excitatory.

A damping term fulfilling these conditions had been specified by Van der Pol (17–19). At this level, other types of damping terms are appropriate as well (e.g., according to the Raleigh differential equation); it is only because of the following supplementations that the Van der Pol type of damping is sufficient exclusively (22). The Van der Pol equation (with the coefficient of frequency taken for unity) reads

$$\ddot{y} + \varepsilon(y^2 - 1)\dot{y} + y = 0 \qquad (2)$$

Solutions of this equation describe self-sustaining and self-excitatory oscillators. When the amplitude (and hence y^2) is small or even zero, the damping is negative, and the amplitude of the oscillation increases; when the amplitude is large, the damping is positive, and the amplitude decreases. After every perturbation, therefore, an intermediate steady-state amplitude adjusts, where the damping is, on the average, zero (21). The character of the resulting oscillation is determined by the "coefficient of damping increment" ε. The case of $\varepsilon \gg 1$ ("relaxation oscillation") is characterized by a large energy dissipation into the environment (by "friction") during parts of the cycle (when the elevation from the zero position is large) and, correspondingly, by a large energy restitution (by the "feedback mechanism") during other parts of the cycle (when the elevation from zero is small). On the other hand, the case of $\varepsilon \ll 1$ ("pendulum oscillation") is characterized by a small energy exchange with the environment. To illustrate the dependence of the oscillation parameters

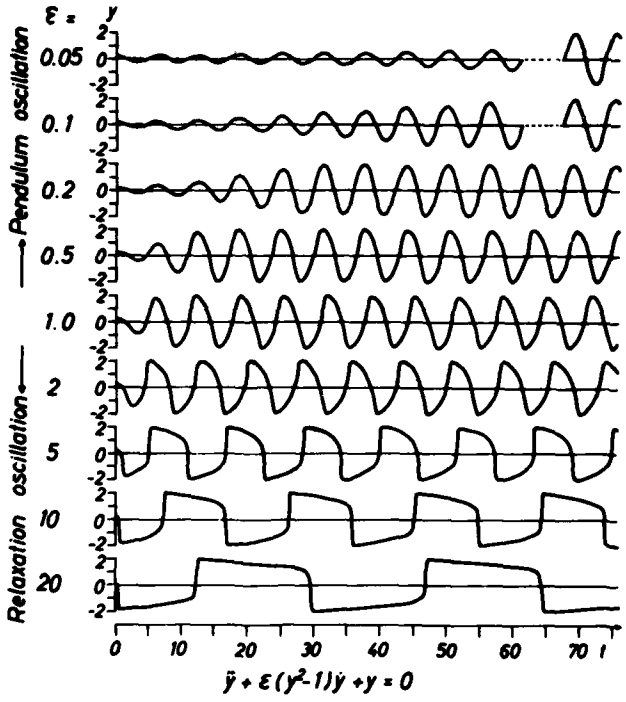


FIG. 1. Solutions of equation 2 with nine different ε values. Initial conditions: $y(0) = 0.2$; $\dot{y}(0) = 0$. (From Wever, ref. 22, with permission.)

on ε, Fig. 1 presents solutions of equation 2 with nine different values of ε. The frequency and waveform of the generated oscillation, as well as the duration of transient processes, depend on ε (22).

When the external force in equation 2 is not zero but is periodically varying, the generated self-sustaining oscillation becomes separately excited, or forced. The frequency of this generated oscillation equals that of the varying environmental stimuli only when the frequency of the forcing oscillation is within a limited range of entrainment. Outside this range the oscillation shows (on the average) its natural frequency (or free-running period). The limitation in the capability to become synchronized is a general property of all self-sustaining oscillations.

To be more general, the "external force" is assumed not to be zero as in equation 2 but z (cf. equation 1). Again, z may be a function of time, e.g., it may vary periodically. The extended Van der Pol equation is then

$$\ddot{y} + \varepsilon(y^2 - 1)\dot{y} + y = z \qquad (3)$$

If z is temporally constant, equation 3 describes self-sustaining oscillations

within a limited range of external forces, or mean values, respectively: $-1 < z < +1$ (21). Inside this "oscillatory range," all parameters of the oscillation depend on z as well as on ε. To illustrate this dependence, Fig. 2 presents solutions of equation 3 computed with the medium value $\varepsilon = 1$ and with nine different z values. As can be seen in this figure, the oscillation loses its symmetry when its mean value deviates from zero; apart from the waveform, the frequency, amplitude, and duration of transient processes depend on z.

With z values outside the oscillatory range, the system remains at rest. When elevated by any disturbance, the oscillation damps out. The return to the original steady state is oscillatory only inside the larger "range of periodic adaptation," $|z| < (1 + 2/\varepsilon)^{1/2}$. With external forces outside the latter range, the system approximates z by a periodic adaptation (22).

In the case of periodically changing z, equation 3 describes an oscillation that is not only self-sustaining but also externally excited; within a limited range of frequencies, it becomes synchronized to the z oscillator. Outside this "range of entrainment," the y oscillator runs on its natural frequency but is rhythmically modified by "relative coordination" with z. The periodic external force affects the oscillation by putting energy into the system; to equalize the energy balance, the amplitude of the forced oscillation enlarges, as does the energy output.

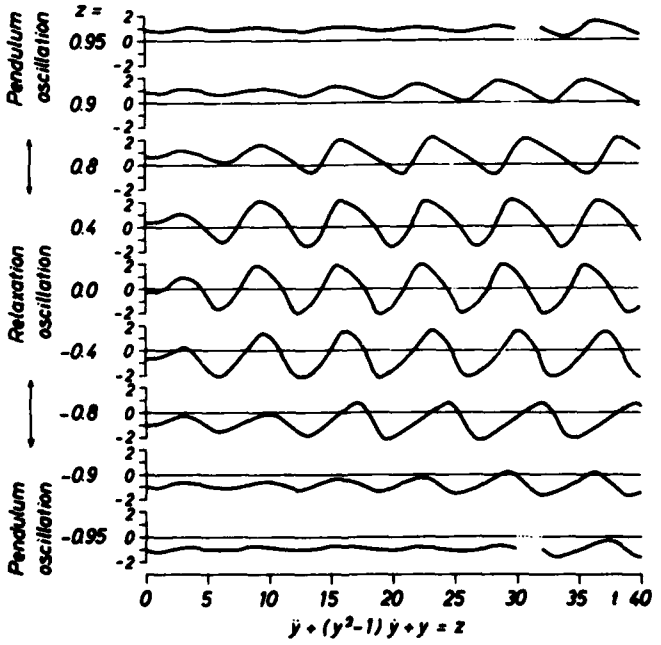


FIG. 2. Solutions of equation 3 with $\varepsilon = 1$ and nine different z values. Initial range of oscillation = 10% of the steady-state range. (From Wever, ref. 27, with permission.)

Consequently, under the influence of a forcing oscillation, the energy exchange between the oscillating system and environment increases, and every oscillation tends toward a relaxation oscillation, in spite of unchanged ε and mean z values. Figure 3 illustrates the superelevation of the amplitude under the influence of external excitation; it shows various solutions of equation 3 computed with and without the influence of a rectangular Zeitgeber. As can be seen, the amount of "resonance superelevation" is greater as ε is smaller and as z deviates from zero.

Solutions of the Van der Pol equation may, in some respects, be appropriate to simulate biological rhythmicity; they are, in other instances, not very well suited to certain biological variables. This is partly due to the fact that the resulting oscillations run either symmetrically around zero (original equation 2, Fig. 1) or at least with changing signs (extended equation 3, Fig. 2). Obviously, when representing biological variables like temperature, concentrations of hormones, or rates in cellular growth, negative values of the variables are meaningless. The absolute value of a variable may be insignificant in a linear oscillation; however, it is of great importance in a nonlinear self-sustaining oscillation where all parameters have been shown to depend considerably on the mean value. Therefore, the model equation must be modified in a manner that excludes alternations in signs:

$$\ddot{y} + \varepsilon(y^2 + y^{-2} - a)\dot{y} + y = z \tag{4}$$

According to this equation, the damping of the resulting oscillation will increase

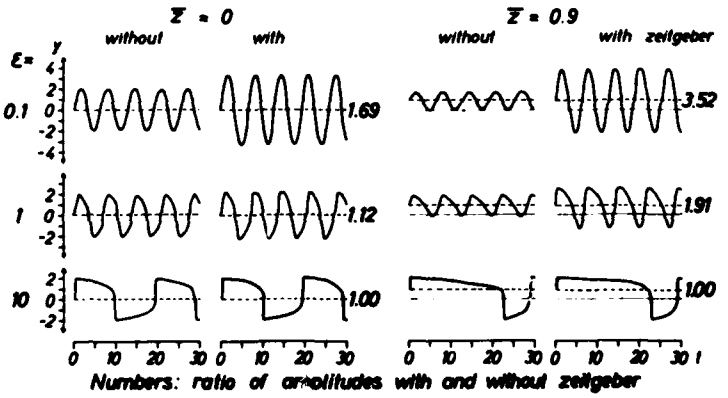


FIG. 3. Steady-state solutions of equation 3 with three different ε values and two different z values; each diagram shows at left the oscillation with constant z value ("autonomous") and at right the oscillation under the influence of a rectangular Zeitgeber with $\Delta z = \pm 0.5$ and a period each coinciding with the corresponding autonomous period ("heteronomous"). *Dotted lines*, mean values. (From Wever, ref. 27, with permission.)

infinitely when y approximates zero, wherein the system would become immovable. The system, therefore, can never transgress zero. In equation 4, a specifies (together with ε and z) the steady-state amplitude of the oscillation, or its "oscillatory strength"; $a < 2$ is a precondition for self-sustainment. If $a = 3$ is set, the minimum value of the net damping is -1, as in the original Van der Pol equation (equation 2); using this value, the oscillatory range is $0.618 < z < 1.618$. To illustrate the effect of the modified damping on the rhythm behavior, Fig. 4 presents solutions of equation 4 in the steady state, with three different ε values and five different z values. Solutions of equation 4 with negative signs are possible as well; then the oscillating variable can never become positive. Consequently, there is another oscillatory range: $-1.618 < z < -0.618$.

In a last step, the value of ε in the nonlinear damping term must be specified numerically, according to biological results. Many independent evaluations meet in the statement that circadian rhythms are positioned close to the middle within the continuum of pendulum–relaxation oscillation; an appropriate average approximation to biological results seems to be $\varepsilon = 0.5$. With this value (and, furthermore, $a = 3$), the range of periodic adaptation is $0.382 < z < 2.618$. In special cases, deviating ε values may better fit biological facts; in these cases, of course, ε may be subject to variations.

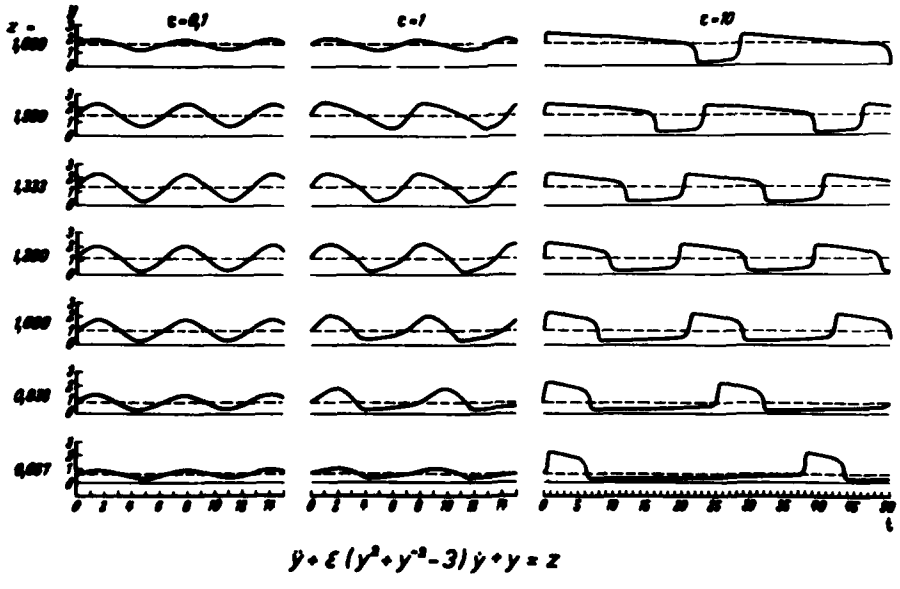


FIG. 4. Steady-state solutions of equation 4 with $a = 3$, three different ε values, and seven different z values. *Dotted lines*, mean values. (From Wever, ref. 22, with permission.)

Nonlinear Restoring Force

In all oscillation equations discussed thus far, the restoring force is linear; that is, the force driving the system back to its neutral position is proportional to the actual elevation from that position. A consequence in the linear equation 1 was the mutual independence of the different rhythm parameters, mainly frequency and mean value. In self-sustaining oscillations (equations 2–4), at least with small ε values, the frequency is independent of the mean value and therefore of the external force. With large ε values, in fact, the frequency varies with the mean value; however, the correlation has an inconsistent sign, even when damping is asymmetrical (cf. equation 4). Conversely, one of the first generalizations derived from biological testing of circadian rhythms in various organisms was that changes in frequency and mean value are consistently positively correlated; this statement constituted the "circadian rule" (1,20). To realize the circadian rule mathematically, a nonlinear restoring force has to be introduced (according to the Duffing differential equation). The least arbitrary way is to replace the frequency coefficient in equation 1 for a power series of y:

$$\omega^2 = 1 + g_1 y + g_2 y^2 + g_3 y^3 + \ldots$$

In a system oscillating around zero, there is a fundamental difference in the meanings of terms with even and odd powers in this series; the resulting restoring force is either symmetric or asymmetric. However, if the oscillation is restricted to values with unchangeable sign, this difference disappears. In this case, therefore, the power series can be broken off after the first term without a relevant loss in specificity. Then the resulting equation reads

$$\ddot{y} + \varepsilon(y^2 + y^{-2} - a)\dot{y} + (1 + gy)y = z \qquad (5)$$

Averaging the results of free-running circadian rhythms in various animal species and in humans, a value of $g = 0.6$ seems to be appropriate to describe the experimentally observed correlation between changes in frequency and mean value. Again, this value is a rough average from various experiments, and in special cases other g values may be more suitable to special biological data. With a nonlinear restoring force, as in equation 5, the external force z is no longer identical with the mean value; with $g = 0.6$, the oscillatory range is $0.847 < z < 3.189$, and the range of periodic adaptation is $0.470 < z < 6.730$.

An alternative to this model may be an oscillation that violates the circadian rule. In such a model, the term describing the restoring force, $(1 + gy)$ in equation 5, must be exchanged for $(1 - gy)$. Such a model, however will not lead to stable oscillations, because the coefficient of frequency will be imaginary, yielding an infinitely increasing or decreasing function (24). Therefore, an oscillation violating the circadian rule cannot be used in this way.

Rather, a precondition for such an application would be a coefficient of frequency that, in fact, decreases with increasing y but can never become negative. A term fulfilling this condition, for instance, would be $\omega^2 = e^{-y}$.

The external excitement by a forcing oscillation is no longer only nonparametric; in addition, it gains a parametric component. It is clear that stable synchronization is possible only if both components of the phase-control mechanism lead to the same phase relationship; this condition is guaranteed only when the circadian rule is fulfilled (20). Even if it seems to be possible to violate the circadian rule in autonomous oscillations, as just discussed, every violation of the rule leads to instabilities in heteronomous rhythms (24). The contribution of a parametric component in separate excitation has, in addition, another consequence. Normally, i.e., in oscillations around zero, parametric excitation leads to periodic solutions with half the period of the forcing oscillations (according to the Mathieu differential equation). However, if the oscillating variable does not change its sign, parametric excitation leads to 1:1 synchronization, like nonparametric excitation. If components of both excitation modes operate in a combined action, the restriction of the oscillating variable as introduced in equation 4 is a precondition for stable synchronization.

The most general expression for the restoring force is the power series mentioned. In fact, if all coefficients g_1 to g_n always equal zero, the result is a stable oscillation with linear restoring force. However, if these coefficients fluctuate randomly in the course of superimposed noise, it may happen that the net value of the nonlinearity becomes negative, and therefore the oscillation becomes unstable (*vide supra*). In order to prevent the system from such instabilities, it is advantageous to set the coefficients mentioned slightly positive in the sense of equation 5; in this case, the oscillation then remains stable even with fluctuating values of the coefficients under consideration.

Another source of stabilization is the continuing mutual interdependence between rhythm parameters effected by the nonlinear restoring force. It effects a negative serial correlation between the durations of successive cycles that stabilizes the frequency of the generated rhythm when exposed to random noise (*vide infra*). Consequently, a nonlinear restoring force, as introduced in equation 5, protects the generated rhythm from disturbing influences of random fluctuations in various respects.

Structure of the External Force

The model equations 3–5 are controlled by an external force. In the modifications of equations 4 and 5, this force must deviate from zero; if the external force varies periodically, constituting a forcing oscillation, its mean value has to deviate from zero. Biological rhythms are, on the average, more or less in phase with the controlling environmental cycles; with a varying ratio between forced and forcing rhythms, the external phase relationship changes its

sign. Depending on the definition of the external stimuli, biological rhythms can also simply run counterphase to the environmental cycles; in this case, also, the external phase relationship changes with varying ratios of the frequencies. Such a phase relationship could be quantified when the external force is not constituted by the environmental stimuli themselves but by their first time derivatives. Then, however, the system would be immovable, because the time derivative is, on the average, zero. Consequently, a combination of z and its time derivatives must be applied.

To achieve mathematical stability, the appropriate combination is that of z with its first and second time derivatives (25). If, to simplify matters, all relevant coefficients are taken for the unit, the resulting model equation reads

$$\ddot{y} + \varepsilon(y^2 + y^{-2} - a)\dot{y} + (1 + gy)y = \ddot{z} + \dot{z} + z \qquad (6)$$

The Resulting Model Equation

The previous sections started with the simple oscillation equation of the second order, and subsequently the two relevant terms were modified by nonlinearities (a coefficient of the term with $\ddot{y}$ can always be abolished by division). The two nonlinearities were compelled by very general summaries of biological experimental results; alternatives to the basic equation, as well as to the modifications, were tested at every step, but all had to be expressly rejected. Finally, the controlling external force was specified with the same method. Through previous deductions, the iterative process of discriminating between alternatives resulted in the model equation 6. For the three free parameters included in this equation, ε, a, and g, numerical values were proposed. Setting these values, the equation constituting the model for circadian rhythmicity reads

$$\ddot{y} + 0.5(y^2 + y^{-2} - 3)\dot{y} + (1 + 0.6y)y = \ddot{z} + \dot{z} + z \qquad (7)$$

In this model equation, the independent variable z represents the external force controlling the system, and the dependent variable y represents the biological variable under consideration. Because equation 7 does not contain any other free coefficient, there is an unambiguous dependence in the behavior of y on z, whether it is constant or varying.

It is evident that solutions of this simple equation 7 can describe only general properties of circadian rhythmicity. Peculiarities of single species, or even of individual organisms, cannot be reproduced by this simple model. Several modifications are necessary when describing properties of a specific rhythm, the first of which involves varying several coefficients. Second, additional terms must be introduced, and the coefficients must be adapted to specific experimental findings. Examples of possible extensions of the models discussed thus far will be given in the following sections.

Systems of Coupled Oscillators

A simplified system of two coupled oscillators can be explained by the following equation:

$$\ddot{y} + \dot{y} + \varepsilon(y^2 - 1)\dot{y} + y = z \tag{8}$$

It depicts an oscillator of the second order capable of self-sustainment and a simple oscillator of the first order that is capable only of damped oscillations. In such a case, it is preferable to speak not of a multioscillator system but of one oscillator generating an oscillation of a more complicated wave shape. Figure 5 presents several solutions of equation 8, computed with various ε and z values. As can be seen, the generated oscillation has, in the case of $\varepsilon \ll 1$, an enlarged amplitude, but it keeps its monomodal shape; the case of $\varepsilon = 1$ results in a bimodal oscillation wherein the relation between the two peaks depends on z; in the case of $\varepsilon \gg 1$, the result is a multimodal oscillation in which the basic oscillation seems to be superimposed by a damped oscillation of a much higher frequency. A great variety of different wave shapes can be generated when, in addition, the term of third order in equation 8 is varied by corresponding coefficients; more details are given elsewhere (22).

It has been mentioned in the context of equation 7 that the external force z

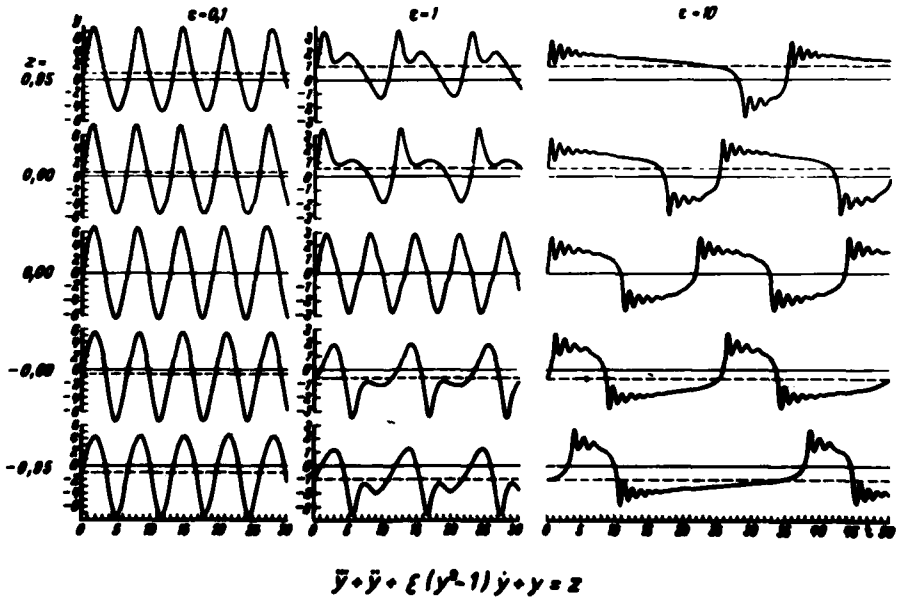


FIG. 5. Steady-state solutions of equation 8 with three different ε values and five different z values. Dotted lines are mean values. (From Wever, ref. 22, with permission.)

represents all those external stimuli that control the generated oscillation. However, it is not cogent that these be solely environmental stimuli; rather, it is possible that they may originate from the output of another oscillator within the same organism (21). This means that the oscillating variable y, generated by one oscillator, is simultaneously a part of the external force controlling another oscillator, x, which, in return, participates in the control of the first oscillator. Starting with equation 6, but with an additional coefficient of frequency according to equation 1, the two-oscillator system is formulated:

$$\begin{aligned} &\ddot{y} + \varepsilon_1(y^2 + y^{-2} - a_1)\dot{y} + \omega_1^2(1 + g_1 y)y \\ &\quad = \omega_1^2[c_1(\ddot{x} + \dot{x} + x) + \ddot{z} + \dot{z} + z] \\ &\ddot{x} + \varepsilon_2(x^2 + x^{-2} - a_2)\dot{x} + \omega_2^2(1 + g_2 x)x \\ &\quad = \omega_2^2[c_2(\ddot{y} + \dot{y} + y) + \ddot{z} + \dot{z} + z] \end{aligned} \tag{9}$$

The only additional coefficients that must be introduced are the coupling coefficients c. Mathematically, this system of two coupled equations of second order each is equivalent to one equation of fourth order. It is arbitrary, therefore, whether solutions of this system are considered as two interacting simple oscillators or one oscillator that is more complex. In the general formulation of equation 9, the environmental input z is fed into the system twice, into the y and the x oscillators; however, the separate control is apparent only because the equivalent fourth-order equation includes only one environmental input (in a more complicated term).

Extension of the Model to Other Frequencies

Complementing the term with y in equation 7 by a frequency coefficient (cf. equation 1) may enlarge the range of applicability of the model equation. With $\omega = 1$, solutions of the model equation have been normalized to result in a period corresponding to about 1 day (*vide infra*). With deviating ω values, they may be applied as models of biological rhythms with other frequencies. The manner in which the meanings of the other terms in equation 7 change with variable ω must be considered.

The external force z keeps its meaning only when multiplied by the same coefficient as the term with y; in particular, the z values defining the oscillatory range keep their numerical values only in this case. The coefficient ε determines the position of the resulting oscillation within the relaxation-oscillation–pendulum-oscillation continuum only with $\omega = 1$; with deviating ω values, this position is determined by ε/ω. This will mean that, with increasing ω, the frequency of the resulting oscillation increases, and, simultaneously, the type of oscillation changes in direction toward a pendulum oscillation. However, evaluations of biological rhythms show that high-frequency rhythms tend more to the relaxation type and low-frequency rhythms more to the pendulum type of oscillation (25). With necessary modification, the equation then becomes

$$\ddot{y} + 0.5\omega^{1.25}(y^2 + y^{-2} - 3)\dot{y} + \omega^2(1 + 0.6y)y = \omega^2(\ddot{z} + \dot{z} + z) \quad (10)$$

This model equation will be applied to describe biological rhythms with high frequencies, particularly the rhythm of the central nervous system (24).

Summary of Preconditions of the Model

The model equation describing circadian rhythmicity is based on a few preconditions:

1. Circadian rhythmicity is based on a feedback mechanism that leads, within a certain range of external conditions, to self-sustaining and self-excitatory oscillations; under all conditions, the oscillating variable does not change its sign.
2. In autonomous rhythms, changes in frequency and mean level are consistently correlated; in heteronomous rhythms, parametric and nonparametric components contribute simultaneously in separate excitations. As a consequence of the intrinsic correlation, the generated oscillation is insensitive to fluctuations in its decisive coefficients.
3. When circadian rhythms are synchronized by external periodicities, the phase-angle difference between biological and environmental rhythms is, where the periods of both rhythms coincide, zero, or the rhythms run counterphased.

These preconditions are sufficient to establish the model equation applicable to single circadian rhythms. If the model should be extended to a system of mutually coupled rhythms, another precondition must be added:

4. If one oscillation influences another, it does so equivalent to the external influence on this oscillator.

If the model should be extended to biological rhythms of very deviating frequencies, e.g., to rhythmic activities of the central nervous system, yet another precondition must be added:

5. If the frequency of the generated rhythm changes considerably because of the introduction of a coefficient of frequency deviating considerably from unity, the position of the resulting oscillation within the continuum of relaxation–pendulum oscillation will change in direction to the relaxation type with increasing frequency.

SOLUTIONS OF THE MODEL EQUATION: PREDICTIONS FOR CIRCADIAN RHYTHMS

In this section, computed solutions of the model equations will be discussed. Because circadian rhythmicity is the main topic, solutions of equation 7 will

be considered first. This equation includes the external force z as the only free parameter; therefore, the solutions to be considered differ only in the environmental conditions controlling the rhythm. If the model is applied to animal rhythms, where light is the most effective environmental stimulus, z can be taken as an analogue to the intensity of illumination. In the version applicable to light-active organisms, z must then be accepted as positively correlated to light intensity (e.g., proportional to the logarithm of light intensity); for nocturnal species, z is negatively correlated to light intensity. If the model is applied to human rhythms, where light is marginally effective, z may be correlated to the amount of social stimuli or to behavioral functions. In any event, every external stimulus affecting circadian rhythmicity is reflected in z.

To enlarge the range of applicability, in what follows we shall consider not only the behavior of the resulting oscillation itself. In this case, y will represent a steadily varying variable, such as body temperature, locomotor activity, or excretion of any substance in the urine. In addition, a square wave will be considered that is derived from the course of the oscillation by introducing a threshold, separating sections where the oscillation runs above threshold ("activity") from sections where it runs below threshold ("rest"); in this case, the derived variable alternating between two discrete states may represent the activity–rest rhythm (20). Moreover, in this case the area between threshold and oscillation during "activity time" may represent the amount of the performed activity. If applied to animal rhythms, where the temporal ratio between activity and rest (α/ρ ratio) is mostly in the range of unity, it seems to be appropriate to set the threshold at $y = 1$. If applied to human rhythms, where the α/ρ ratio is commonly in the range of two, the threshold has to be set lower, at about $y = 0.5$. Finally, solutions of equation 7 have periods in the range of four units of time. Because this period should correspond to the circadian period, all solutions of equation 7 are normalized so that the unit of time corresponds to 6 hr, of four units of time correspond to 24 hr.

Autonomous Rhythms

Rhythms under Constant Conditions

When the external conditions are considered to be constant, only z itself has a finite value, whereas $\dot{z}$ and $\ddot{z}$ are constantly zero. Under this condition, equation 7 describes a remaining oscillation only as long as z is within the limited oscillatory range, $0.847 < z < 3.189$. Figure 6 presents solutions of equation 7 with seven different constant z values, all within the oscillatory range; from every solution, three steady-state cycles are drawn. With increasing z value, the period shortens, the amplitude transgresses a maximum, and the wave shape changes from skewed to the right ("form factor" < 1) to skewed to the left (form factor > 1). The separation of activity and rest by a threshold constitutes

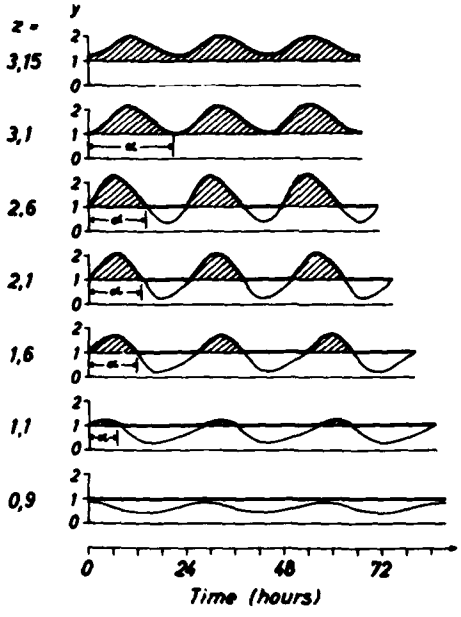


FIG. 6. Steady-state solutions of the model equation 7 with seven different z values; from every solution, three successive cycles are drawn. A threshold at $y = 1$ separates the course of the oscillation in sections above the threshold ("activity" α; *hatched areas*) and below the threshold ("rest" ρ; *open areas*). (Adapted from Wever, ref. 29.)

additional rhythm parameters. Only in the two most extreme diagrams, such a separation is not possible, though the rhythmicity persists. In the lowermost diagram the persisting rhythm remains below threshold, describing "continuous rest," and in the uppermost diagram the similarly persisting rhythm remains above threshold, describing "continuous activity." In the remaining diagrams, the activity time α lengthens with increasing z, despite the shortening period. The amount of the performed activity increases over the full range with increasing z (25). Figure 6 does not show that the duration of transient states also depends on z, being shortest in about the middle of the oscillatory range.

Outside the oscillatory range, the system remains at rest. Only when pushed by any stimulus does the system move, but it returns to its constant steady state. Inside the range of periodic adaptation, i.e., with $0.470 < z < 0.847$ and $3.189 < z < 6.730$, the feedback mechanism is sufficient to compensate for friction to such a degree that the system returns to its steady-state value of oscillation. In other words, inside these two ranges adjacent to the oscillatory range, the system is capable of damped oscillations. Only outside these ranges, i.e., with $0 < z < 0.470$ and $z > 6.730$, is the net damping so large that the system adapts to its steady state aperiodically.

Rhythms Under the Influence of Random Noise

The assumption of the deterministic model, as applied in the preceding section, was not realistic; it was presented only to make obvious the

interdependences between different rhythm parameters. In nature, the observed rhythms are always disarranged by superimposed random fluctuations. Therefore, of special interest is the computation of solutions of the model equation under the influence of random noise. All following presentations are based on random fluctuations of the external force. Preliminary computations show that the results are similar when other coefficients of equation 7 fluctuate randomly.

Figure 7 shows two solutions of equation 7, computed with different mean values of the external force z but with equal random fluctuations of z. From both solutions, we present not only the generated rhythms themselves but also the square-wave rhythms derived by using the threshold that separates activity α and rest ρ; for once, the threshold is taken at $y = 0.5$ to describe rhythms with an α/ρ ratio generally larger than unity, as is the case in human rhythms. As can be seen in this figure, not only the means of the period and amplitude are different in the two rhythms (Fig. 6) but also the variabilities of these two parameters. Generally, the rhythm with the longer period and the smaller amplitude is much more sensitive to standard random disturbances, or it is less precise (38).

Successive cycles within the time series, when fluctuating because of random noise, are not independent of each other. Rather, there are serial correlations between corresponding phases in successive cycles. At first, the durations of successive cycles are negatively serially correlated; this holds true with the maximum and minimum values for reference, as well as with "activity onset" and "activity end" for reference, i.e., with the passings of the threshold. This internal stabilization causes an accumulation of the cycle-to-cycle variations that is much smaller than it would be with randomly distributed deviations of the

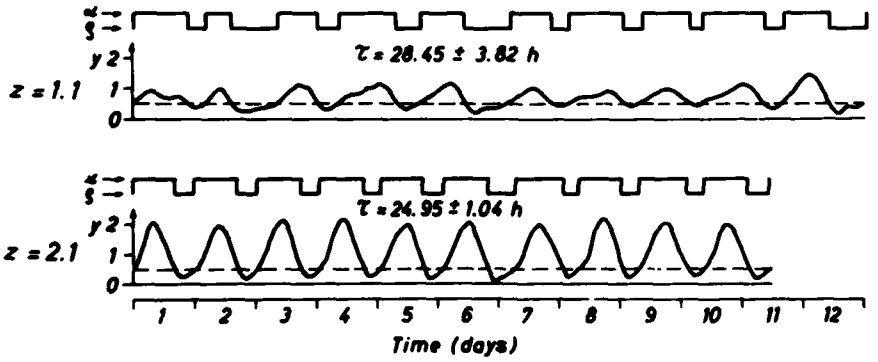


FIG. 7. Two solutions of the model equation 7 with different mean z values and superimposed random fluctuations of z with equal variabilities; from every solution, 10 successive cycles are drawn. In addition to the courses of the rhythms, the derived square waves (with the threshold at $y = 0.5$) are drawn. (From Wever, ref. 38, with permission.)

same magnitude; in other words, it causes a ratio between stability and precision that increases much less with increasing duration of the time series than with a random series (56). In addition, there is another negative serial correlation between activity time and rest time. This correlation can be observed between an activity time and the following rest time, but also with a slightly smaller probability between a rest time and the subsequent activity time. Consequently, the stabilization mechanism just mentioned operates on the full cycle, and it does not stabilize the separated sections within a cycle (36).

In particular, the serial correlation between successive cycles deserves attention, because such an internal stabilization mechanism is commonly attributed exclusively to a coupled system consisting of "pacemaker" and "overt rhythm" (*vide supra*). In equation 7, however, it is an inherent property of the isolated system, due to the nonlinear restoring force. This becomes obvious in Fig. 8, where two different solutions of the model equation are presented, with a linear restoring force (left) and a nonlinear restoring force (right). In all respects except the coefficient of nonlinearity in the restoring force, the two solutions are equivalent; in particular, they are computed with equivalent superimposed random noises. They are only normalized to equal

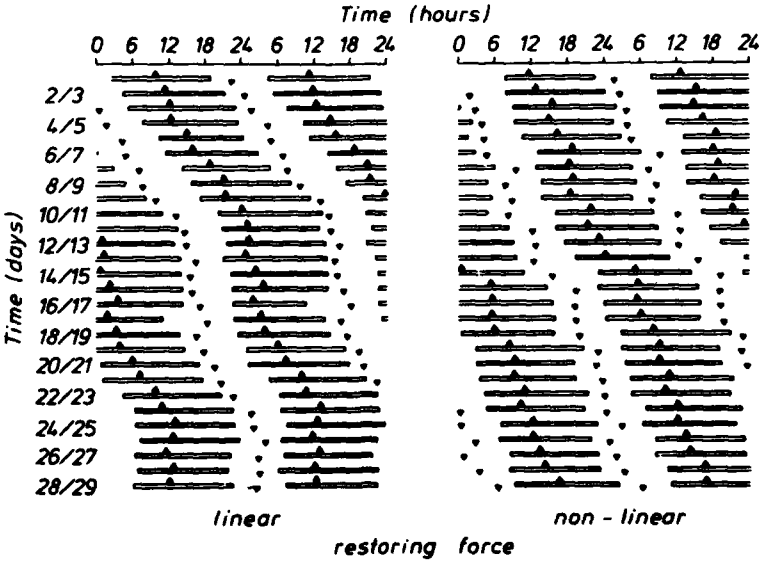


FIG. 8. Two solutions of the model equation 7 under the influence of the same random noise (random fluctuations of z), however with linear restoring force ($g = 0$) (*left*) and normalized to the same period as the solution with the nonlinear restoring force (*right*). The rhythms are presented in the manner of biological data (double plots); "activity" is represented by bars, and from the course of the rhythm only the temporal positions of maximum (*open triangles*) and minimum values (*filled triangles*) are given. Successive cycles are drawn one beneath the other.

periods (without the normalization, the period with the linear restoring force would be slightly longer than that with the nonlinear). The solutions are presented in the same manner in which human rhythms are frequently presented: The rhythm's courses are separated into "activity" and "rest" by a fixed threshold, and the maxima and minima of the rhythms are indicated. As can be seen, the two solutions behave differently with regard to long-term fluctuations: The model with the nonlinear restoring force is, in the long run, much more stable than the model with the linear restoring force, despite the slightly larger cycle-to-cycle variations. This stability is due to the negative serial correlation between the durations of successive cycles ($r_s = -0.409$) in the model with nonlinear restoring force; the correlation is zero in the other model.

The stabilization mechanism mentioned is based on the intrinsic coupling between changes in the actual values of frequency and amplitude. The coupling that is due to the nonlinear restoring force compensates for the effects of perturbations on the phase of the rhythm. Therefore, the stabilization mechanism can be effective only in the range around $\varepsilon = 1$ (i.e., the middle between pendulum and relaxation oscillations), where both frequency and amplitude are subjected to changes following any perturbation (42). Neither in pendulum oscillations ($\varepsilon \ll 1$), where the amplitude is changed by a perturbation but not the frequency, nor in relaxation oscillations ($\varepsilon \gg 1$), where the frequency is changed but not the amplitude, can the intrinsic coupling be effective. Therefore, the demand for long-term stability of the generated oscillation restricts the coefficient of damping increment to values close to $\varepsilon = 1$. It is remarkable that just these values are also demanded by the results of biological experiments.

The solutions of the model equations shown in Fig. 8 are accidental realizations out of a great variety of possibilities. However, all the different realizations show the same essential results: Solutions with a nonlinear restoring force consistently show a negative serial correlation, whereas solutions with a linear restoring force show a serial correlation close to zero. The accidental differences in the patterns of the generated rhythms reveal a pitfall in the analysis of circadian time series. Comparing the first 10 cycles and the second 10 cycles (as frequently performed in biological analyses) in the solution with a linear restoring force (Fig. 8, left), the mean cycle durations differ, and the short-term, or cycle-to-cycle, variations result in standard deviations to guarantee statistically significant differences in periods. Of course, such a statement is completely meaningless, because two accidental samples from the same basic entity are considered; it is based on an inadequate application of statistics neglecting fundamental preconditions. Because of mutual interdependence, successive cycles within a time series cannot be used for statistical purposes, although such a consideration may be of a high descriptive value; a "period" is defined only by considering a total time series consisting of many successive periodically repeated intervals, not by considering single intervals. Of course, a correct period analysis of any type does

not allow differentiation between the periods of the first and second 10 days of this example.

In the contrast to the serial correlation among successive cycles, that between adjacent sections within a cycle is independent of the structure of the restoring force. With linear as well as nonlinear restoring force, an activity time and the following rest time are negatively correlated, as are (but to a lesser degree) a rest time and the following activity time. The level of these negative serial correlations depends on the slope of the rhythm at the point where it crosses the threshold. In the relevant range, this slope is flatter during the descending part in the rhythmic course of the variable, where the threshold crossing marks "end of activity" or "onset of rest," than during the ascending part. Consequently, the variability of "end of activity" is larger than that of "onset of activity," and the serial correlation between an activity time and the following rest time is more negative than that between a rest time and the following activity time.

Interdependence of Rhythm Parameters

In preceding sections, the dependences of several rhythm parameters on the external force have been discussed. In fact, the same correlation between isolated rhythm parameters and certain external conditions can be realized with various types of model equations if the free coefficients are selected in a proper way. Therefore, agreement between model predictions and experimental data yields very little information about the structure of the system as long as it is restricted to a few points. Of relevance is simultaneous experimental confirmation of many diverse predictive modes. Therefore, each model must determine as many different rhythm parameters as possible and also the mutual interdependence between all these parameters.

Figure 9 summarizes, necessarily incompletely, the dependences of various rhythm parameters on z as calculated from equation 7; several of these parameters have already been mentioned with the discrete solutions of this equation as presented in Figs. 6–8. It is characteristic that many of these parameters are positively correlated to z within the largest part of the oscillatory range; simultaneously, this means that within this range the parameters mentioned are positively correlated with each other. It is only close to the limits of this range that opposite correlations occur. In the following, the parameters presented will be discussed separately.

The range of oscillation, or the rhythm's amplitude, deviates from zero only inside the oscillatory range; its maximum is not in the middle of this range, but close to its upper end, at $z = 2.58$. The sensitiveness of the amplitude against a standard noise is not least when the amplitude is largest, but at a lower z value. Even if the power of the superimposed noise is not constant but is a fixed percentage of the amplitude, the variability is least at about the middle of the oscillatory range, and it increases toward the ends. The course of the mean value had been presupposed in the modeling process; it is always positively

correlated to z and therefore cannot be rated as a prediction. The positive correlation of the rhythm's frequency to the mean value is the substance of the "circadian rule" (*vide supra*); it is worth noting that the course of frequency passes a weakly marked minimum close to the lower end of the oscillatory range. The sensitivity of this frequency to the standard noise is least at about the z value where the amplitude is largest, i.e., at a distinctly higher z value than the sensitivity of the amplitude. If the power of noise is a fixed percentage of the respective rhythm amplitude, its influence on the variability of the frequency is nearly independent of z, again in contrast to that of the amplitude. The difference in the sensitivities of amplitude and frequency against standard noise is especially obvious in the diagram showing the "precisions" (ratios between means and standard deviations) of these parameters depending on z.

Apart from the common rhythm parameters, several more parameters can be considered and are therefore included in Fig. 9. The next parameters are the "form factor" (ratio between descending and ascending parts of the cycle) and the "ratio of deviations" (between the variations of end and onset of activity); these two parameters are closely related. With small z values, the shape of the rhythm is skewed to the right; under the influence of random noise, the onset of activity varies more than does the end of activity. With higher z values, the shape of the rhythm is skewed to the left, and the end of activity is more affected by superimposed random noise than is the onset of activity. Close to the limits of the oscillatory range, both parameters tend toward unity (like other parameters defined as ratios, both of these parameters are drawn on logarithmic scales).

The introduction of a threshold separating activity time and rest time creates additional rhythm parameters. The ratio between activity time and rest time (α/ρ ratio) increases, with common threshold at $y = 1$, continuously with increasing z; it approaches zero close to the lower limit and approaches an infinite value close to the upper limit of the oscillatory range (Fig. 6). With alternative thresholds, however, the course of the α/ρ ratio deviates. With a lower threshold (i.e., with generally high α/ρ ratios) it becomes negatively correlated to z when z is small: with a higher threshold (i.e., with generally low α/ρ ratios) it becomes negatively correlated to z when z is large. The amount of activity increases steadily with increasing z; because this statement is valid with all thresholds, the amount of activity is presented only with the medium threshold at $y = 1$.

Particularly interesting is the course of the "internal phase-angle difference" between the oscillation itself and its square-wave derivative. It is evident that the "true" phase-angle difference can only be zero, independent of z, because the two rhythms represent different manifestations of the same oscillation. Nevertheless, the diagram shows that the formally computed phase-angle difference between these two manifestations is not constantly zero (positive phase-angle differences mean that the acrophase of the complete rhythm leads that of its square-wave derivative). With the "standard threshold" ($y = 1$), the

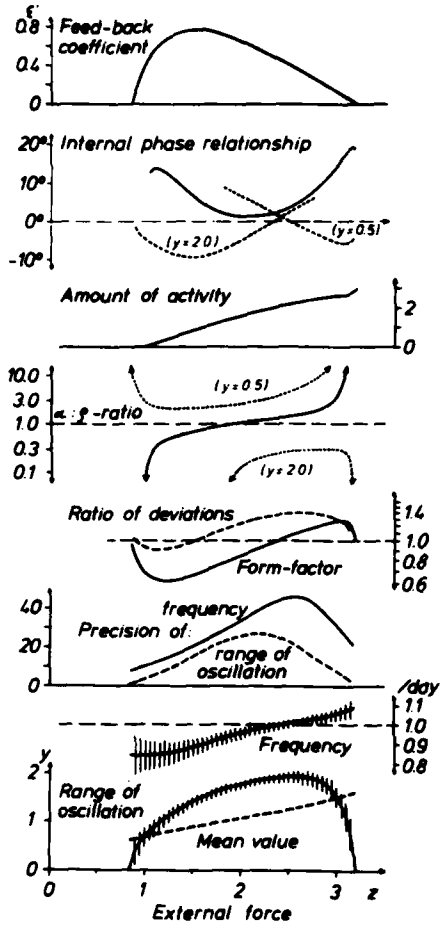


FIG. 9. Solutions of the model equation 7: several rhythm parameters depending on the external force z. From range of oscillation and frequency, means and formally computed standard deviations of successive cycles are presented, calculated under the influence of "standard random noise." The other rhythm parameters are explained in the text.

phase-angle difference is always positive, but it varies with z. With the higher and lower thresholds, in fact, the phase-angle differences show even changing signs depending on z, although the courses of the correlations are different in the two cases. Inconsistencies in formal computations are clearly understood to be due to the combined changes in wave shape and in the α/ρ ratio. This inconsistency should be a warning about extrapolating from such a formal computation (which, in principle, is based on sine waves) to "true" phase-angle differences without reference to a special model (43).

Finally, the uppermost diagram of Fig. 9 illustrates the "energy exchange" per cycle between an oscillating system and the environment, expressed in the coefficient ε'; it is calculated according to previous models (25). This diagram demonstrates that the model oscillation as derived from equation 7 tends more to the relaxation type near the middle of the oscillatory range and more to the

pendulum type toward the limits of this range; this result is in agreement with Fig. 2. It is remarkable that this analytical estimation of the feedback coefficient from equation 7 coincides with an empirical estimation based on the behavior of the oscillation under the influence of random noise: It has been shown that the coefficient of damping increment equals the ratio between the coefficients of variability (reciprocal value of precision) in period and amplitude (42). In a pendulum oscillation ($\varepsilon' \ll 1$), the amplitude is relatively more altered than is the frequency by superimposed distortions, and in a relaxation oscillation ($\varepsilon' \gg 1$), the frequency is relatively more altered than the amplitude.

Rhythms under Special Conditions

In addition to the conditions discussed thus far, there are other states in which rhythms run autonomously, but in which the external conditions are neither constant nor randomly fluctuating (*vide supra*) but instead are regularly changing.

In experiments excluding the natural day–night cycle, it may be the case that the external conditions are not really constant but are self-controlled by the activity–rest cycle of the experimental organism. An example is a self-controlled light–dark cycle in which it is light during the activity time and dark during the rest time. Of course, other external stimuli (e.g., ambient temperature) can be self-controlled as well. Beyond this, the hypothesis has been offered that the self-control mechanism can also be behavioral. This is, for instance, the case if an organism in constant illumination has open eyes during activity time and closed eyes during rest time and if the eyes have been shown to be the pathway in the entrainment of the circadian system by the light (35,44).

A self-controlled cycle of environmental stimuli necessarily lags behind the controlling biological rhythm; i.e., light can be switched on only after the onset of activity, not before. Because of the strong correlation between period and phase relationship (*vide infra*), this lag slows down the period of the controlling rhythm, until a shifted equilibrium between period and phase relationship has been adjusted. Consequently, according to equation 7, self-control always lengthens the autonomous period of a free-running rhythm.

The period-lengthening effect of this self-control is stronger where the amplitude of environmental stimuli is larger (e.g., the larger the difference in the intensities of illumination in a light–dark alternation). Because lower values of z are frequently limited in their potential to vary (e.g., total darkness), the mean value of z usually increases with the amplitude. Under "constant conditions," the period will shorten with increasing z (cf. Fig. 6). There are, therefore, two opposite effects: the continuous-action approach, in which the correlation between z and the period is negative; the self-control effect, in which the correlation is positive. These effects may neutralize each other, more or less. The self-control effect concerns only the period, without affecting the other rhythm parameters like mean value. Therefore, changes in these parameters hold their original correlations to changes in the external force also in the

presence of self-control. Although fulfillment of the circadian rule is a precondition for the general stability of the oscillating system (*vide supra*), this inconsistency may explain why this rule seems to be violated in some organisms. However, this inconsistency is valid only with diurnal behavior; with nocturnal behavior, the continuous-action approach is reversed, but the self-control effect is not, so that both effects operate in the same direction. If the self-control effect should be evaluated separately, without being obscured by the effects of possible changes in the mean of z, solutions of an equation with a linear restoring force (cf. equation 3 or 4) should be computed.

Another state in which rhythms run autonomously under regularly changing conditions is realized in an experimental environment in which the light intensity (or the ambient temperature) is slowly but consistently changing (51). Whereas under constant conditions the time derivatives of z in equation 7 are constantly zero, under steadily changing conditions $\dot{z}$ is no longer constant but has a value deviating from zero. In a self-sustaining oscillation, the limits of its oscillatory range are defined by the sum of z and its time derivatives. Consequently, the limits of the oscillatory range under steadily increasing z (i.e., with a positive value of $\dot{z}$) are reached with a lower absolute z value than under constant conditions (where $\dot{z} = 0$); they are reached with a higher absolute z value with steadily decreasing z values. Therefore, because of the effects of the time derivatives of the external force, the influence of steady changes in the external force on the system will be opposite to that assumed when hysteresis is presupposed (51).

Heteronomous Rhythms

The number of possible predictions increases considerably when equation 7 is computed with a periodically changing external force. Already, in autonomous rhythms, in which the external force has only one degree of freedom in which it can vary, numerous different predictions have been derived from the solutions of the model equation. In heteronomous rhythms, the external force has several more degrees of freedom in which it can vary. Therefore, the variety of possible predictions to be made is difficult to survey. Within the scope of this brief summary, only some samples of predictions can be presented.

Influence of Different Zeitgeber Parameters

In this section, variations in the temporal course of a periodically changing external force will be considered, i.e., the dependence of the forced rhythm on the shape of the forcing oscillation, or Zeitgeber. In accordance with most experimental studies of circadian rhythmicity, a trapezoidal shape of the Zeitgeber is presupposed, i.e., the external force alternates, with interposed "twilight" transitions, between two fixed states. With such a Zeitgeber, five different parameters can be varied: (a) the difference (or the ratio) between the

values of the external force during the two states; (b) the period of the alternation between the two states; (c) the mean level of the external force; (d) the temporal ratio ("LD ratio") between the two states; (e) the duration of the twilight transitions between the two states. Effects of variations in the respective parameters of the Zeitgeber will be considered in this sequence. To keep the number of solutions to be presented within reason, each of the other Zeitgeber parameters is kept constant at a medium value (or at two values, in case of the range of the Zeitgeber): Unless explicitly stated otherwise, the range of the Zeitgeber is set at $\Delta z = \pm 0.1$ and ± 0.5, the period is $T = 24$ hr, the mean level is $\bar{z} = 2.1$ [when $z = 2.1$ is set continuously, the autonomous rhythm has a period of 24.9 hr (cf. Fig. 6)], the LD ratio is 1:1, and the twilight duration is 5% of the period (1 hr 12 min in the 24-hr day). The conclusions to be drawn from the different solutions are restricted, in this section, primarily to the phase relationship between the forced rhythm and its forcing Zeitgeber (25).

Figure 10 illustrates the influence of the range of the Zeitgeber; it indicates the influence of the Zeitgeber's strength. In this and the following figures, the strip below the rhythm indicates the course of the Zeitgeber. White means higher z values ("light"), and black means lower z values ("dark"); the transitions between white and black indicate the twilights. Figure 10 demonstrates that, apart from drastic changes in the wave shape, the phase of the rhythm advances relative to that of the Zeitgeber with increasing Zeitgeber strength. In order to keep the energy input affected by the Zeitgeber comparable to that affected by the feedback mechanism (*vide supra*), the two strongest of the five different Zeitgeber strengths presented will not be used in the following.

Figure 11 illustrates the influence of the Zeitgeber period. With the weak Zeitgeber (left), synchronization takes place only for periods between about 22 and 28 hr (25 ± 3 hr); secondary ranges of entrainment of about half this width

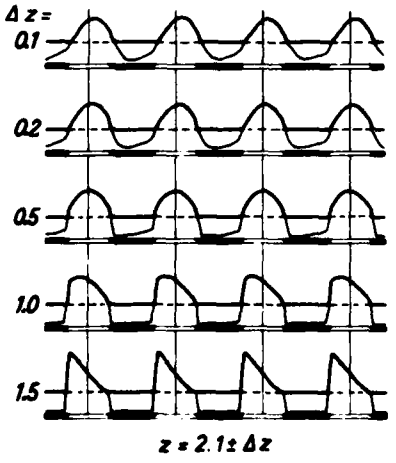


FIG. 10. Solutions of the model equation 7 under the influence of a Zeitgeber, with five different ranges of the forcing (trapezoidal) oscillation. (From Wever, ref. 25, with permission.)

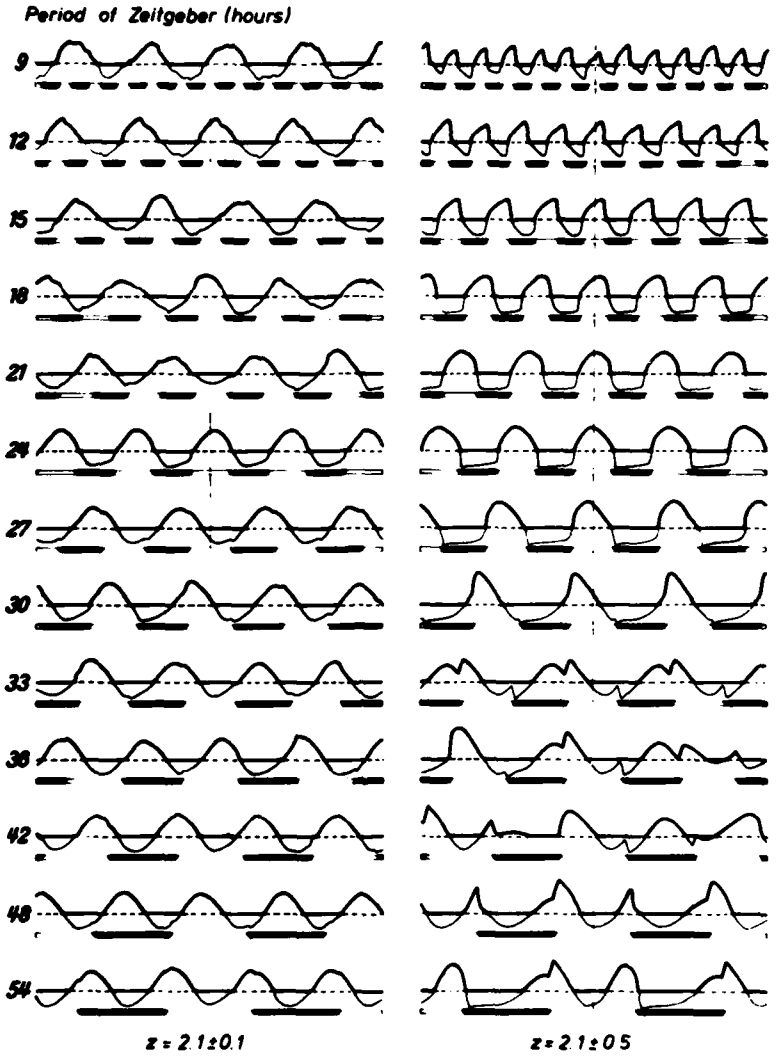


FIG. 11. Solutions of the model equation 7 under the influence of a Zeitgeber, with 13 different periods and two different ranges of the forcing (trapezoidal) oscillation. The ranges of periods where the rhythms are synchronized ("ranges of entrainment") are indicated by vertical lines. (From Wever, ref. 25, with permission.)

are positioned around 12.5 hr (synchronization to half the Zeitgeber period) and around 50 hr (synchronization to twice the Zeitgeber period). With all other Zeitgeber periods, the rhythm free-runs but is affected by the Zeitgeber ("relative coordination") (*vide infra*). With the strong Zeitgeber (right), the ranges of entrainment are about three times broader. The range of periods seems

to be unlimited to short periods, at first glance; however, closer inspection shows that the rhythm is only apparently synchronized to the shortest periods (*vide infra*). Within the ranges of the entrainment (in Fig. 11 indicated by vertical lines), the phase of the rhythm advances relative to that of the Zeitgeber with lengthening period, and the amplitude is largest in the middle of these ranges.

Most evident in Fig. 11 are the limitations of the ranges of entrainment. However, it must be considered that the statement of an unambiguous size of this range is a simplification. In nonlinear oscillations, one must differentiate between a larger "range of holding" and a smaller "range of catching" a rhythm (21). When the period of a Zeitgeber that previously synchronized a rhythm is slowly changed, the limits of entrainment are expanded, depending on the changing speed. When, on the other hand, a previously free-running rhythm is exposed to a Zeitgeber, or the period of a Zeitgeber that previously did not synchronize the rhythm is changed in the direction of the intrinsic period, the Zeitgeber is capable of starting synchronization of the rhythm only within considerably narrower limits. This behavior may resemble "learning," although it is but a mathematical consequence of nonlinearity (27).

Figure 12 illustrates the influence of the level of the Zeitgeber. It shows that the phase of the rhythm advances relative to that of the Zeitgeber when the level increases. This change in the external phase relationship is more obviously marked with the weaker Zeitgeber than with the stronger Zeitgeber; with the stronger Zeitgeber, changes in its intensity are primarily expressed in variations

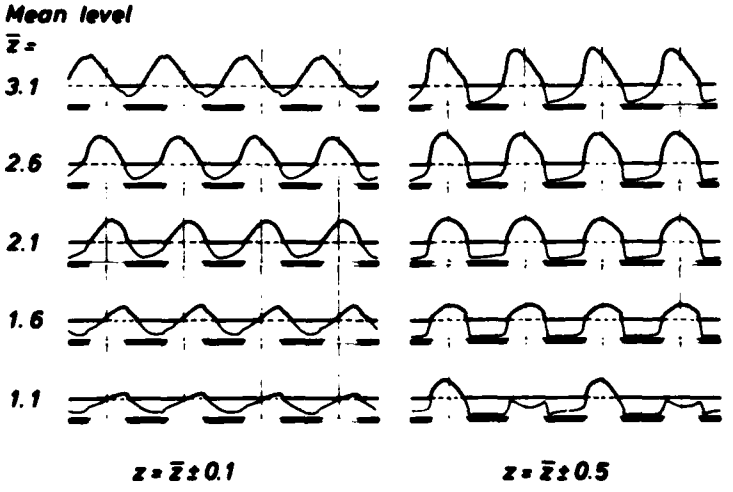


FIG. 12. Solutions of the model equation 7 under the influence of a Zeitgeber, with five different mean levels of the forcing (trapezoidal) oscillation. (From Wever, ref. 25, with permission.)

of the wave shape. With the lowest Zeitgeber level, a special phenomenon is indicated: When a variable with an already low level is influenced by a strong signal, lowering it even more, it can be "frozen" for a while. It does not return to the original wave shape with the next upward push, but only with the second; consequently, activity occurs only during every second light time.

Figure 13 illustrates the influence of the LD ratio. With the weak Zeitgeber, the phase of the rhythm relative to that of the Zeitgeber is most nearly coinciding when the Zeitgeber is symmetric (LD = 12:12), and it delays with both increasing and decreasing LD ratios. This change in the external phase relationship is mainly due to a change in the strength of the Zeitgeber with the LD ratio (the Zeitgeber is strongest when symmetric). With a strong Zeitgeber, there is a consistent advancing of the rhythm's phase relative to that of the Zeitgeber when the LD ratio increases. This change in the external phase relationship is due to the preponderance of changes in the mean level of the Zeitgeber with changing LD ratio over changes in the Zeitgeber strength (the mean level is $\bar{z} = 2.1$ only with a symmetric Zeitgeber; it varies between $\bar{z} = 2.48$ with LD = 21:3 hr and $\bar{z} = 1.72$ with LD = 3:21 hr when the Zeitgeber is strong, but only between $\bar{z} = 2.18$ and 2.02 when the Zeitgeber is

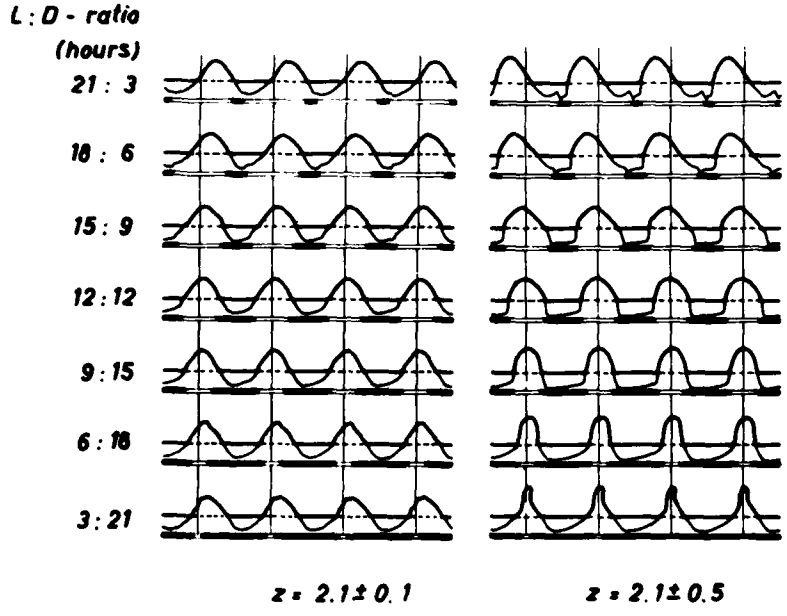


FIG. 13. *Solutions of the model equation 7 under the influence of a Zeitgeber, with seven different LD ratios (temporal ratios between sections with higher and lower z values) and two different ranges of the forcing (trapezoidal) oscillation. (From Wever, ref. 25, with permission.)*

weak). Again, with a strong Zeitgeber, variations in wave shape with changing LD ratios are considerable.

Finally, Fig. 14 illustrates the influence of twilight duration on the rhythm, computed with three different LD ratios but only with the stronger Zeitgeber ($\Delta z = \pm 0.5$). Independent of the LD ratio, the phase of the rhythm advances relative to that of the Zeitgeber with increasing twilight duration. In addition, the twilight duration has a remarkable influence on the dependence of the external phase relationship on the LD ratio. With the shortest twilight, the phase relationship is nearly independent of the LD ratio; conversely, there is strong dependence with the longest twilight—the larger the LD ratio, the earlier the rhythm's phase. Therefore, in addition to this advancing effect, it is suggested that a lengthening of the twilight duration operates like a strengthening of the Zeitgeber (25).

The influences of the various Zeitgeber parameters on the phase relationship between rhythm and Zeitgeber can be summarized by a few generalizations: (a) A stronger Zeitgeber yields a closer phase relationship; conversely, with a weaker Zeitgeber the absolute value of the phase-angle difference between rhythm and Zeitgeber is larger, independent of its direction. (b) A shorter intrinsic period of the rhythm relative to the Zeitgeber will phase-advance its period relative to that of the Zeitgeber; the rhythm's intrinsic period is determined by the level of the external force (*vide supra*). (c) With lengthening of the twilight duration, the relative weights of z and its time derivatives (*vide supra*) are shifted so that a phase advance results.

When speaking of light and dark, usually we are referring to "light-active" behavior (*vide supra*): The external force is positively correlated to the intensity of illumination. When "dark-active" behavior is to be considered, several of the analogues to biological behavior are reversed. This especially concerns the

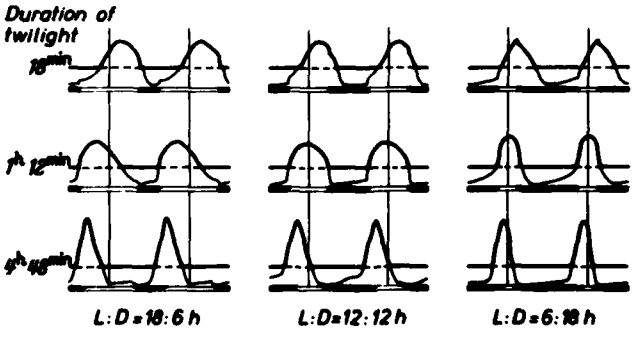


FIG. 14. Solutions of the model equation 7 under the influence of a Zeitgeber, with three different "durations of twilight" (sections of transitions between constantly higher and lower z values) and three different LD ratios of the forcing (trapezoidal) oscillation. (From Wever, ref. 25, with permission.)

influence of the LD ratio: In nocturnal organisms, the phase of the rhythm should be earlier the smaller this ratio (provided the twilights are sufficiently long). This reversion, however, does not concern the influence of the twilight duration. In the course of nature's seasons, there are variations in not only the LD ratio but also the twilight duration; the course of the external phase relationship (depending on the season), as computed for nocturnal behavior, is therefore not simply a reflection of that for diurnal behavior. On the basis of data for LD ratio and twilight duration from 45 degrees of latitude, computations for diurnal behavior result in a marked early phase in midsummer and a consistently late phase for nearly half a year during winter. Computations for nocturnal behavior result in two moderate but nearly equal phase advances in midsummer and midwinter and, correspondingly, late phases in spring and fall (30).

The Masking Effect

Under the influence of a forcing oscillation, or a Zeitgeber, with varying parameters, the rhythm changes its wave shape. This is especially obvious in Figs. 10–14 with a stronger Zeitgeber. These changes in wave shape are due in part to a phenomenon that is caused by the cooperation of time derivatives of the external force (*vide supra*). If solutions of equation 7 are considered where the rhythm and Zeitgeber are considerably out of phase, it suggests that the rhythm would be composed of two components: an endogenous component of a shape not unlike the sinusoidal shape and an exogenous component that runs directly parallel to the Zeitgeber. Of course, the solution of equation 7 describes a homogeneous rhythm, and the two components are inherently inseparable. However, with a very rough approximation, the solution of equation 7 can be separated into two components by splitting the equation:

$$\ddot{y}_1 + 0.5(y_1^2 + y_1^{-2} - 3)\dot{y}_1 + (1 + 0.6y_1)y_1 = \dot{z} + \bar{z} \tag{7a}$$

$$y_2 + 0.6y_2 = z - \bar{z} \tag{7b}$$

In these equations, $\bar{z}$ is the mean level of the external force. In this system, y_1 as the solution of equation 7a will describe the endogenous component, and y_2 in the solution of equation 7b will describe the exogenous component: $y = y_1 + y_2$ then describes the combined rhythm.

A "masking effect" has frequently been observed in circadian experiments (1,49). This term describes a course of the observed variable (e.g., locomotor activity) that reflects directly changes in the Zeitgeber (e.g., light intensity). For instance, in a diurnal organism, introducing light during the organism's natural nighttime can evoke a burst in locomotor activity at a phase where the organism should be at rest; similarly, enforcing darkness during its "day" can block locomotor activity. This masking effect is especially obvious when the

locomotor-activity rhythm and the light–dark cycle have different periods; it then tends to superimpose itself over the organism's "true" clock rhythm.

It is a reasonable assumption to set the apparent superimposition of endogenous and exogenous rhythm components described by equation 7 as analogous to the apparent combination of the "true" rhythm with a masking effect. In fact, the similarities are striking. If the biological rhythm is considered as being composed of two components (with and without using the clock), instead of being a uniform system, the mathematical equivalent will be the system of equations 7a and 7b. The solutions of equation 7 and the sums of the solutions of equations 7a and 7b are similar, but they differ in several details. For instance, the feedback is controlled by y in equation 7 but by y_1 in equation 7a, i.e., only by one of the two components constituting y. Moreover, the contribution of the second time derivative of the external force is missed in equation 7a and, hence, the influence of the twilight duration on the rhythm (Fig. 14) is different in the two alternatives. More important, according to equation 7 the degree of the masking effect depends on the movability of the system, which is a function of its elevation. The movability is highest with medium elevations, where the "friction" is smallest (*vide supra*); consequently, the masking effect is largest with medium elevations, i.e., around the turning points of the oscillation. The movability (and hence the masking effect) is smallest around the maximum and minimum values of the oscillation. On the other hand, according to the combined system described by equations 7a and 7b, the amount of the masking effect is independent of the phase of the rhythm where it occurs.

The masking effect, according to experimental evidence, seems to be phase-dependent, with maximum values around the turning points and minimum values around the extremes of the oscillation (6). This means that the masking effect actually seems to be an inherent property of the "clock system," therefore using the same physiological pathways as the rhythm itself. The alternative explanation—a separate pathway of the masking effect bypassing the clock, as expressed in equations 7a and 7b—must be rejected. The masking effect can therefore be stated as the direct and obvious reflection of a Zeitgeber. Certainly the masking effect is by no means restricted to self-sustaining oscillations; it is also present outside the oscillatory range where the system is not self-sustaining but is capable of exerting forced oscillations.

Borderline Cases of Entrainment

Self-sustaining oscillations can be synchronized by external periodicities only within limited ranges of entrainment. Inside these ranges they are synchronized to a forcing oscillation; outside these ranges they maintain their own endogenous period, in spite of the presence of the forcing oscillation. Consequently, the transition from inside to outside the oscillatory range results in a

considerable change in the measurable period of the rhythm. The change in period, however, should not be considered an abrupt alteration between two constant period values; rather, close to the limits of the range of entrainment, special phenomena occur that manifest themselves in the regular modulations in phase and amplitude of the rhythm. Because these phenomena are general properties of all self-sustaining rhythms, they can be verified with the original Van der Pol equation; to simplify matters, therefore, the following demonstrations are calculated as solutions of equation 2.

Outside the range of entrainment, the energy exchange between Zeitgeber and oscillating system (*vide supra*) is not sufficient to synchronize their periodicities. This interaction leads to relative coordination (12) that is more pronounced as the period approaches the limit of the range of entrainment. Relative coordination means that, in fact, all phase-angle differences between rhythm and Zeitgeber occur successively but are not equally distributed. When the period is closest to that of the Zeitgeber, the amplitude of the rhythm is maximal; it is minimal at the opposite phase, when the period deviates most from the Zeitgeber period. Figure 15 presents two examples of relative coordination. The left diagram shows the solution of equation 2 with $\varepsilon = 0.2$, with a Zeitgeber period slightly shorter than the lower limit of the range of entrainment; the right diagram shows a rhythm with $\varepsilon = 5$ under the influence of a Zeitgeber whose period is slightly longer than the upper entrainment limit. Figure 15 shows that the phenomenon of relative coordination is independent of the type of the self-sustaining rhythm (39).

Inside the range of entrainment, but close to its limits, a regular modulation of period and amplitude can occur. Figure 16 shows two examples of "relative entrainment" (39). Here, the phase-angle difference between rhythm and Zeitgeber fluctuates, but covering only a part of the full cycle; in the left diagram, no less than 180° are covered, and in the right diagram 65°. Again, the Zeitgeber period is, in the left diagram, close to the lower limit and, in the right diagram, close to the upper limit of the range of entrainment. In both examples the rhythm seems to free-run for several cycles, but with continuous reduction in its amplitude; then it returns to its original phase within one or two cycles, while the amplitude increases again. In the left diagram, the amplitude is reduced to such a degree that it is difficult to differentiate whether the phase jumps forward or backward, or to distinguish between relative entrainment and relative coordination; in the example shown at the right, however, there is no doubt about its classification. Finally, there are borderline cases between relative and absolute entrainment. The right diagram of Fig. 17 shows, for comparison, the normal type of absolute entrainment; after a phase shift of the Zeitgeber, the phase of the rhythm adjusts to its original phase aperiodically, or asymptotically. In the left diagram the rhythm adjusts, after the same phase shift, periodically fluctuating to its steady-state phase, i.e., the phase adapts in the course of a damped oscillation. There are important differences between these two examples; at the right, the period of the Zeitgeber is close to the

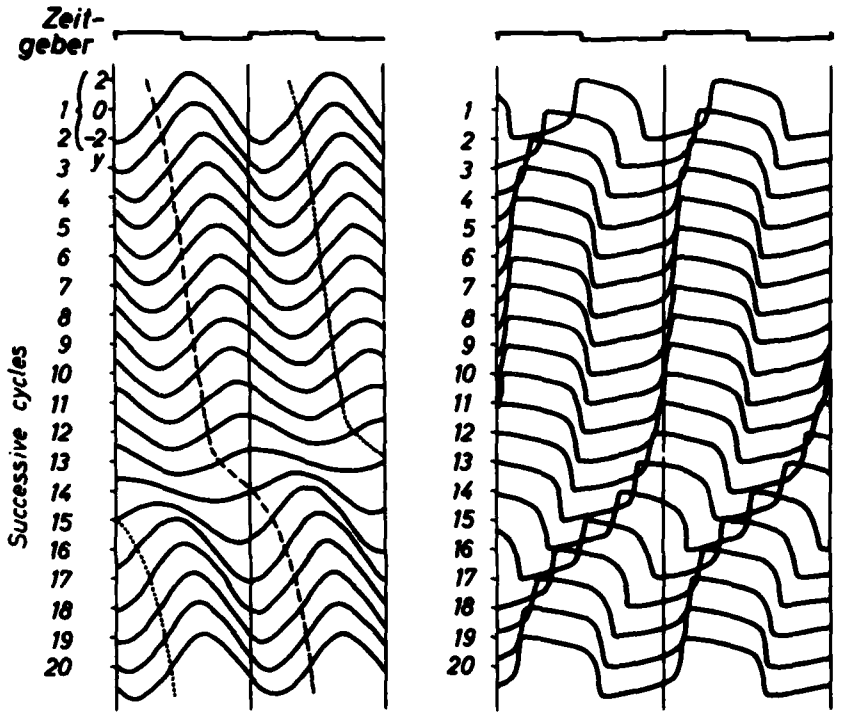


FIG. 15. Two solutions of equation 2 under the influence of a rectangular forcing Zeitgeber, with periods just outside of the range of entrainment. **Left:** $\varepsilon = 0.2$; period of the Zeitgeber shorter than the lower entrainment limit. **Right:** $\varepsilon = 5$; period of the Zeitgeber longer than the upper entrainment limit. The dotted line combines corresponding phases of successive cycles. (From Wever, ref. 39, with permission. Copyright: Academic Press Inc., London, Ltd.)

middle of the range of entrainment, whereas at the left, it is close to the upper limit of this range. It is another consequence of this difference in the Zeitgeber period that the final phase relationship between generated rhythm and Zeitgeber differs, in the two diagrams, about 90°.

The borderline cases of entrainment, as shown in Figs. 15–17, can be observed in all self-sustaining oscillations; for instance, they are present in solutions of equation 7 as well as those of the simple equation 2. Therefore, the mere existence of the phenomena just described tells little about the underlying equation except that it describes self-sustaining oscillations. There is only a weak dependence on the coefficient ε: The larger ε, the smaller is the range of Zeitgeber periods where these phenomena occur. Moreover, the same phenomena occur when a limit of the range of entrainment is transgressed not because of a change in the Zeitgeber period (or, with fixed Zeitgeber period, because of a change in the intrinsic period of the oscillation) but because of a change in the

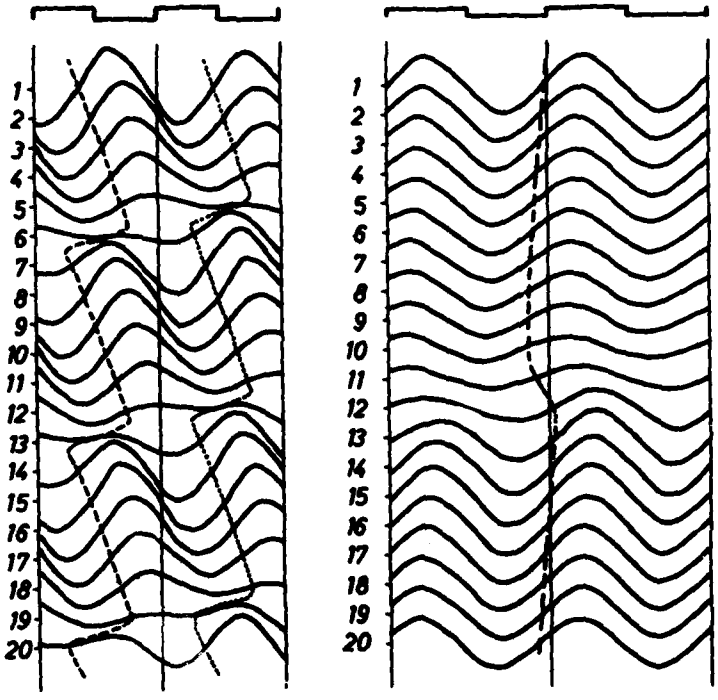


FIG. 16. Two solutions of equation 2 under the influence of a rectangular forcing Zeitgeber, with periods just inside of the range of entrainment; $\varepsilon = 0.2$. Left: Period very close to the lower entrainment limit. Right: Period close to the upper entrainment limit. The dotted lines combine corresponding phases of successive cycles. (From Wever, ref. 39, with permission. Copyright: Academic Press Inc., London, Ltd.)

amplitude of the Zeitgeber, or its "strength"; also in this case, the limits of the range of entrainment shift relative to the period of the oscillation. Further details of the borderline cases mentioned are given elsewhere (39).

Influence of Self-Sustainment

Phenomena near the limits of the range of entrainment have been considered where self-sustainment of the rhythms has been a precondition; only self-sustaining rhythms have finite limits of entrainment. On the other hand, rules derived from the numerous solutions with varying Zeitgeber parameters are valid, as well with rhythms that are not self-sustaining but are capable of damped oscillations. In fact, many properties as observed under the influence of a Zeitgeber are very similar to self-sustaining and non-self-sustaining rhythms; this is true, for instance, for the three rules determining the rhythm's phase (*vide*

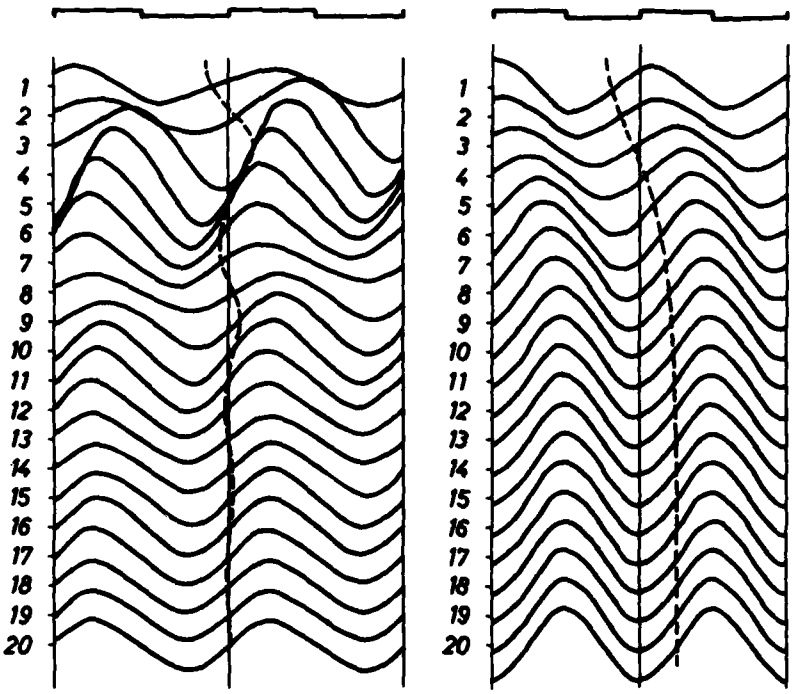


FIG. 17. Two solutions of equation 2 under the influence of a rectangular forcing Zeitgeber, with periods inside of the range of entrainment; $\varepsilon = 0.2$. **Left:** Period close to the upper entrainment limit. **Right:** Period in the middle of the range of entrainment. In both diagrams, the Zeitgeber had been phase-shifted for 90° at day 0. The dotted lines combine corresponding phases of successive cycles. (From Wever, ref. 39, with permission. Copyright: Academic Press Inc., London, Ltd.)

supra). The precondition for fulfilling the mentioned rules is that the system be within the range of periodic adaptation. Outside this range, a system can react only passively to changes in external forces; consequently, in heteronomous rhythms the larger range of periodic adaptation is more important than the smaller oscillatory range. In autonomous rhythms, however, the latter range is more important, because it is only within this oscillatory range that the system maintains a self-sustaining oscillation, whereas it comes to rest outside this range.

Figure 18 shows the behavior of a rhythm (as the solution of equation 7) under the influence of a wide range of external forces. On the left, the behavior under the influence of a Zeitgeber is shown. As can be seen, there are steady transitions in phase relationship, amplitude, and wave shape as long as the system is inside the range of periodic adaptation; only after leaving this range (uppermost diagram) does the behavior alter considerably. At the right side of

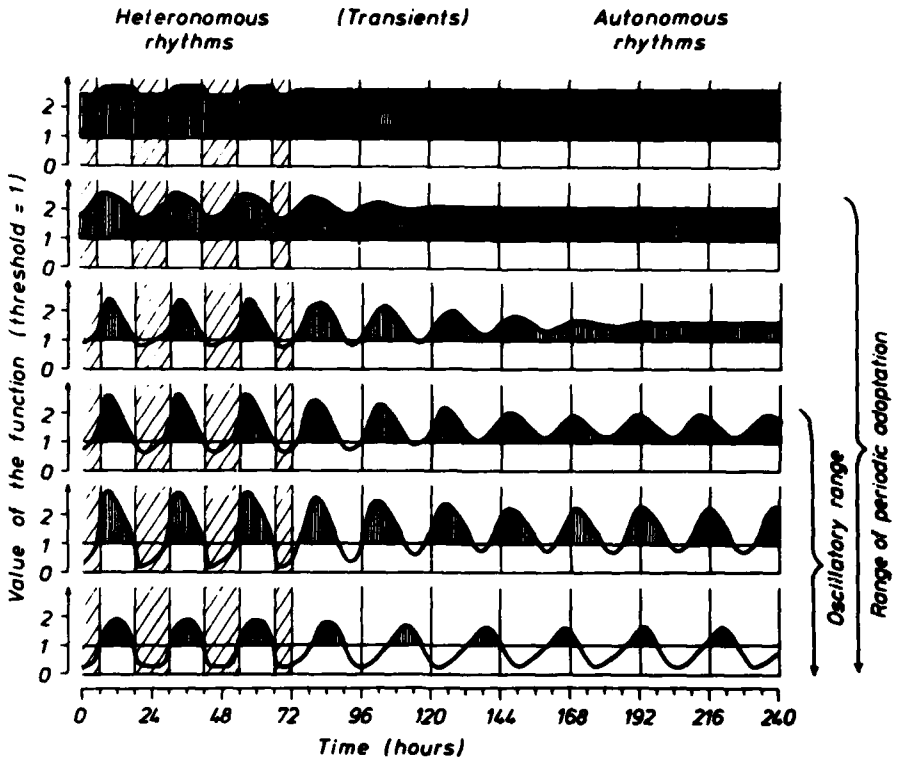


FIG. 18. Solutions of the model equation 7 with six different external forces z (increasing from bottom to top), at left under the influence of a 24-hr Zeitgeber ($\Delta z = \pm 0.5$) and at right under constant external forces, with transient states in between. At the right border, the revelant ranges of z values are indicated. (From Wever, ref. 50, with permission.)

Fig. 18, the behavior of the same rhythm under constant conditions is shown. Evidently, the range of periodic adaptation is without relevance (at least, after the transients), and only the oscillatory range is important. In between both heteronomous and autonomous rhythm states, transient states are demonstrated; they last longest as the rhythm approaches the limits of the oscillatory range. Of course, Fig. 18 covers only a part of the possible external forces z. With further decreasing external forces, the other limit of the oscillatory range (and then that of the range of periodic adaptation) is transgressed until, eventually, with very low z values, a range of aperiodic reactions to external stimuli is again reached.

It is sometimes overlooked that most circadian experiments are performed within a range of experimental conditions that is relatively small in comparison with the range of external conditions present in the natural environment. This is partly due to the fact that in larger ranges of constant experimental conditions

(e.g., when the intensity of illumination is too high or the ambient temperature is too low), most organisms become arrhythmic—conditions that are frequently stated to be "insufficient" for circadian experiments. However, it has been shown that even under the latter conditions, circadian rhythms are present when the conditions vary periodically, in which case all properties of the synchronized rhythms correspond to properties observable under conditions where the rhythms free-run (50). As shown in Fig. 18, these conditions must correspond to external forces inside the range of periodic adaptation but outside the oscillatory range.

Considering only the behavior of the system inside the smaller oscillatory range, the circadian system has been described as an "oscillator." Consideration of the whole range of conditions where circadian phenomena can be observed, which includes the larger range of periodic adaptation, leads rather to the description of it as a "resonance amplifier" (50). In fact, this amplifier becomes "overmodulated" in a certain range of conditions (the oscillatory range) where it initiates a self-sustaining oscillation; however, it holds its amplification capacity within a much larger range. The mathematical basis of this system is the active reduction of the "net damping" by a feedback mechanism, expressed in the negative value of a in equation 6, or the term "-3" in equation 7. Within certain ranges of external forces this mechanism overcompensates the positive damping due to "friction," with the result of self-sustainment. In larger ranges, however, the feedback mechanism reduces the net damping to the extent that the system remains capable of oscillating under the influence of periodically varying external conditions, in spite of relatively high "friction." Without the feedback mechanism in operation (i.e., with $a = 0$), the system would react periodically only in a relatively small range (and even then only with ε values smaller than unity).

Zeitgeber Phase Shifts

In the previously presented solutions of equation 7 describing heteronomous rhythms, the Zeitgeber was always temporally constant, and the rhythm was always shown in the steady state; only in the borderline cases of entrainment was there no steady state, but rather regular fluctuations. However, it is of special interest to see in what way a steady state is reached—specifically, to look for the transients. A particularly appropriate and reliable way to do so is to look for the behavior of the rhythm following a phase shift of the Zeitgeber (28), which may also be of practical interest (53). In the following, therefore, solutions of equation 7 are presented based on a phase shift of the Zeitgeber; in all computations, the mean level of the Zeitgeber is $\bar{z} = 2.1$, and unless otherwise stated, its range is $\Delta z = \pm 0.2$.

At first, solutions with phase shifts of the synchronizing Zeitgeber for 6 hr are computed. Figure 19 shows some different possibilities. Columns A and C show 6-hr delays, and columns B and D show 6-hr advances; in columns A and B

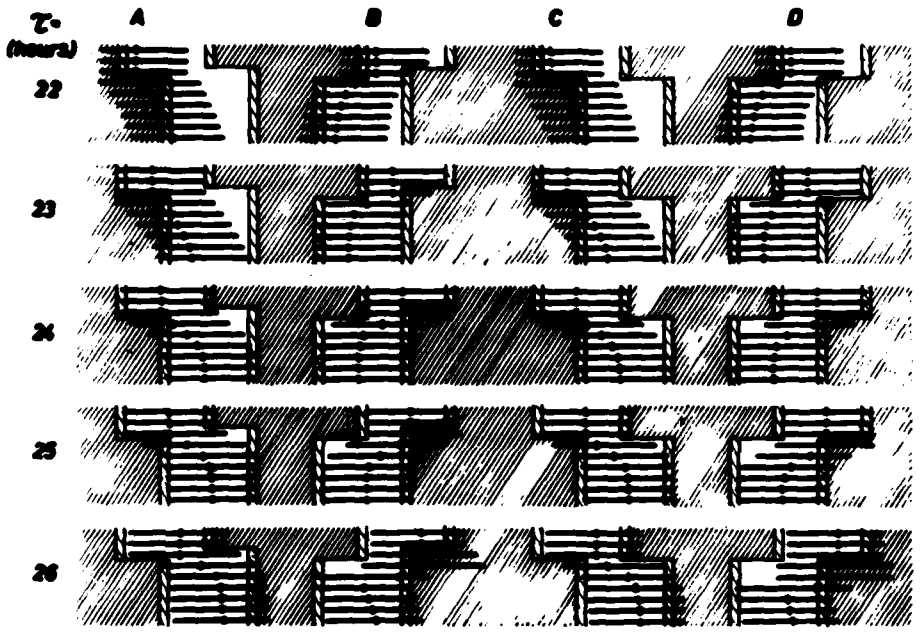


FIG. 19. Solutions of the model equation 7 under the influence of a 24-hr Zeitgeber, the phase of which is shifted for 6 hr at the third day. In particular, equation 7 is modified by a "frequency coefficient," resulting in five different intrinsic periods τ. The phase of the Zeitgeber is shifted by lengthening a single "light time" for 6 hr (*column A*) or shortening it for 6 hr (*column B*), or by lengthening a single "dark time" for 6 hr (*column C*) or shortening it for 6 hr (*column D*). Successive "activity episodes" are indicated by bars, drawn one beneath the other. Marks at the bars: maximum values of the oscillation. (From Wever, ref. 28, with permission. Copyright: Academic Press Inc., London, Ltd.)

once a "light time" had been altered, and in columns C and D once a "dark time." Deviating from the previous computations, a coefficient of frequency is added in equation 7 and set so that the autonomous rhythms have periods of exactly 22, 23, 24, 25, and 26 hr; it is therefore possible to study the influence of the intrinsic period τ on the reentrainment behavior (the Zeitgeber always has a period of 24.0 hr). In Fig. 19, the bars indicate activity time (oscillating variable above the threshold at $y = 1$), and the marks at the bars indicate the maxima of the rhythms. At first glance, reentrainment is not symmetric following delays and advances. With the 24-hr rhythm, the duration of reentrainment is clearly shorter after 6-hr advances than after 6-hr delays, in spite of a coincidence in the periods of rhythm and Zeitgeber; only with the 25-hr rhythm are the durations equivalent. The delaying phase shifts are always answered by regular reentrainment: The duration of reentrainment decreases steadily with increasing intrinsic period of the rhythm. In contrast, reentrain-

ment after advancing phase shifts is not so regular: The duration of reentrainment is shortest with medium intrinsic periods, and it increases with changes in the period of the rhythm; with the slowest rhythm, advancing shifts are answered by an overshooting reentrainment. When the intrinsic period is still slower ($\tau = 27$ hr), the 6-hr advancing Zeitgeber shifts are accompanied by 18-hr delays of the rhythm; delaying Zeitgeber shifts does not result in comparable behavior change. Finally, reentrainment is expedited in all cases when the Zeitgeber shift is released by a single alteration of a dark time instead of a light time.

To describe the irregular reentrainment behavior after advancing Zeitgeber shifts, Fig. 20 presents longitudinally the courses of the rhythms from column B in Fig. 19. Here, particularly with the 26-hr rhythm, the amplitude is drastically reduced for some cycles following the Zeitgeber shift. It has been shown that a temporary reduction of the amplitude below its steady-state value speeds up the rhythm during the transient state (22). This indicates that the duration of reentrainment is shortened after advancing shifts and lengthened after delaying phase shifts when the amplitude of the rhythm is reduced during reentrainment (53). In general, a closer inspection of the reentrainment with regard to wave shape shows that during reentrainment the rhythms are significantly altered. The result is, in general, faster reentrainment of "end of activity" than of "activity onset" and faster reentrainment of the minima than of the maxima.

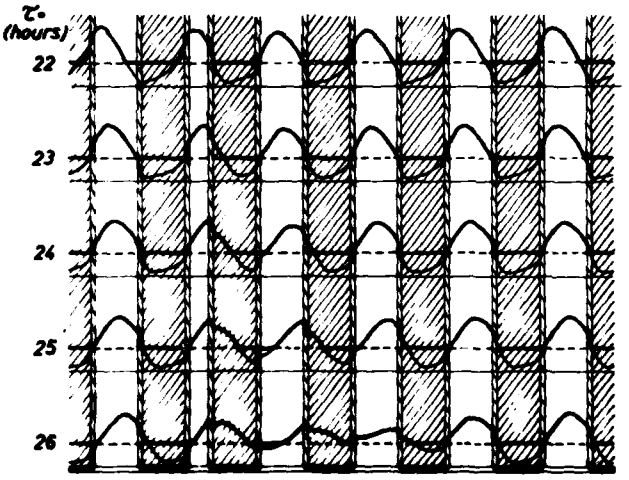


FIG. 20. Solutions of the model equation 7 under the influence of a 24-hr Zeitgeber, the phase of which is advanced for 6 hr at the second day by shortening a single "light time" for 6 hr (cf. column B in Fig. 19). In particular, equation 7 is modified by a "frequency coefficient," resulting in five different intrinsic periods τ. The courses of the oscillations are presented for six successive cycles. (From Wever, ref. 28, with permission. Copyright: Academic Press Inc., London, Ltd.)

Zeitgeber shifts of other varying degrees have also been examined. With advancing Zeitgeber shifts, the probability of reentrainment via the longer pathway is greater the larger the Zeitgeber shift. Consequently, 12-hr phase shifts of the Zeitgeber are always responded to by delays of the 24-hr rhythm, whether the Zeitgeber shift occurs by doubling a light time or a dark time (Fig. 21). The same Zeitgeber shifts are always responded to by phase advances of the 22-hr rhythm; with the 23-hr rhythm, the amplitude is reduced after the 12-hr Zeitgeber shift to such an amount that the interpretation is ambiguous.

All computations discussed thus far have been computed with the same Zeitgeber of medium strength; there remains the influence of the strength of the Zeitgeber. Figure 21 shows two solutions with a 24-hr rhythm and a 12-hr phase shift of the Zeitgeber (by a doubling of light time) with a common Zeitgeber and a Zeitgeber 2.5 times stronger. As can be seen, reentrainment is obviously faster with the stronger Zeitgeber than with the weaker Zeitgeber. Also, in the diagram with the stronger Zeitgeber the influence of the "masking effect" (*vide supra*) is particularly obvious. Shortly before the 12-hr-delayed "lights off" the system becomes active by passing the threshold; the immediately following "lights off" forces down the variable by masking (and hence blocking the activity) to such an extent that it transgresses the threshold again only about 6 hr later. Similar results concerning the influence of the strength of the Zeitgeber can be observed experimentally with all other Zeitgeber shifts. More details of these computations are given elsewhere (28).

Influence of Random Fluctuations

The deterministic model is not realistic for heteronomous rhythms, just as little as for autonomous rhythms; circadian rhythms are, in addition, influenced by random noise when under the influence of a synchronizing Zeitgeber. However, with heteronomous rhythms, the number of meaningful predictions is

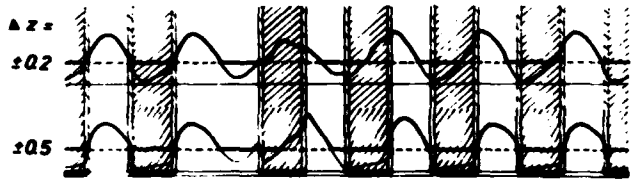


FIG. 21. Solutions of the model equation 7 under the influence of a 24-hr Zeitgeber, the phase of which is delayed for 12 hr at the second day by doubling a single "light time." In particular, equation 7 is modified by a "frequency coefficient," so that the intrinsic period is 24 hr. In the two solutions presented, the strengths or the ranges of the Zeitgeber are different. The courses of the oscillations are presented for six successive cycles. (From Wever, ref. 28, with permission. Copyright: Academic Press Inc., London, Ltd.)

already great when refraining from the additional influence of such stimuli. In the interest of brevity, only some of the possible indications will be given; they concern effects where the influence of superimposed random noise deviates from that with autonomous rhythms (*vide supra*) and where it qualitatively changes the properties of heteronomous rhythms, as discussed in the previous sections.

In heteronomous rhythms, long-term stability is controlled by the Zeitgeber (which has, at least in laboratory experiments, no variability in itself). Those rhythms show, therefore, a negative serial correlation. It is for this reason that serial correlations tell little, if anything, about the mechanisms of generating the rhythms as they do in autonomous rhythms (*vide supra*).

A rhythm that is synchronized to a Zeitgeber with a period close to an entrainment limit and is superimposed, in addition, by random fluctuations may transgress this limit spontaneously because of an accidental elevation of excessive amount. The result is a rhythm that is no longer synchronized to the Zeitgeber. In this case (leaving the range of entrainment), the entrainment limit is defined by the larger "range of holding" (*vide supra*). However, for spontaneous resynchronization (reentering the range of entrainment), the smaller "range of catching" will be applicable; here the entrainment limit for spontaneous resynchronization is shifted closer to the intrinsic period of the rhythm. Therefore, spontaneous resynchronization would need another accidental elevation of a larger amount and would therefore be more improbable than spontaneous desynchronization. In summary, it is probable that a previously synchronized rhythm spontaneously loses its synchrony to the Zeitgeber rather than becoming spontaneously resynchronized by the same Zeitgeber.

The range of periods where spontaneously external desynchronization can occur as a consequence of random fluctuations is similar to that range where the borderline cases of entrainment occur (*vide supra*). This means that the latter range can be overcome when random fluctuations are especially large; in this range, the borderline cases cannot occur, and they are therefore restricted to rhythms with relatively small variability. Because the range where the borderline cases can occur becomes smaller with increasing values of ε (*vide supra*) the probability for the occurrence of the borderline cases decreases with increasing ε values of the constituting equation.

Phase Response Curves

Solutions of equation 7 have been computed describing undisturbed autonomous and heteronomous rhythms. Of interest also is the behavior of an autonomously running rhythm after exposure to single perturbations. Every temporary change in the external force, or every "stimulus" that synchronizes the rhythm when given periodically, is capable of phase-shifting the rhythm. It is the expression of the phase-control mechanism on which the synchronization is

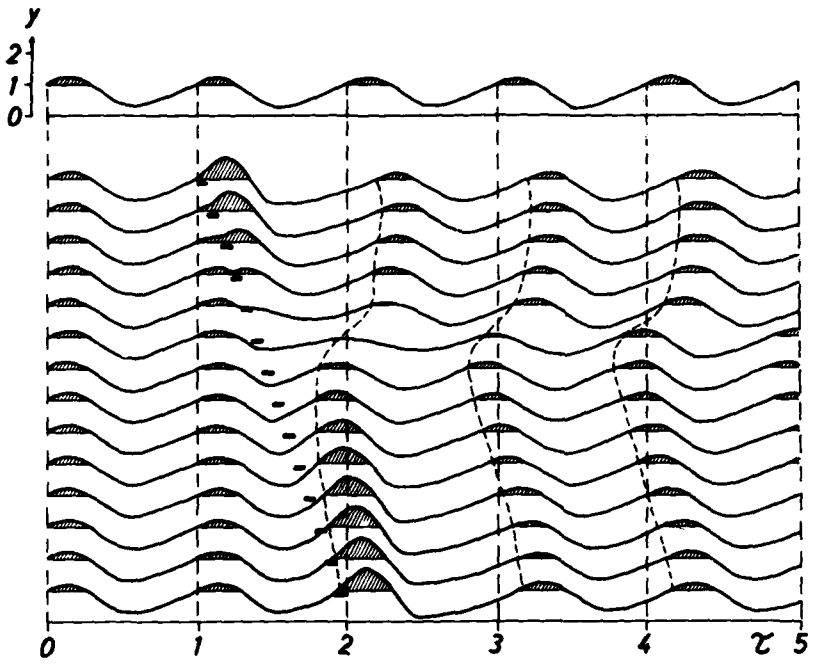


FIG. 22. Solutions of the model equation 7 under the influence of solitary "light pulses" (increase of z for $\Delta z = 2$, lasting for 2 hr). In the upper diagram, the undisturbed oscillation with $z = 1.1$ is presented for comparison. In the lower diagrams, the courses of the oscillation are presented with the single "light pulse" at 14 different phases of the oscillation. The dotted lines combine "activity onset" in the different oscillations that are differently phase-shifted by the single perturbations ("phase response curves"). (From Wever, ref. 25, with permission.)

based that a phase shift released by the stimulus depends in amount (and mostly also in direction) on the phase of the rhythm hit by the stimulus (20). In circadian experiments, phase response curves have indicated that the released phase shifts are a function of the phase of the releasing stimulus. In the following, corresponding computations based on equation 7 are presented.

Figure 22 demonstrates the generation of a phase response curve. The same autonomous rhythm ($z = 1.1$) is pushed by a solitary increase of z ($\Delta z = 2$) lasting 2 hr, repeated successively 14 times at different phases. As can be seen, the single perturbation causes, at several phases, a delay of the rhythm and, at other phases, an advance; simultaneously, the amplitude of the rhythm is partly increased and partly decreased because of the stimulus. If the stimulus hits the rhythm shortly before activity onset, the first cycle following the stimulus is advanced, but after the transients fade, the rhythm is delayed relative to the control (uppermost diagram). As a rule, when the stimulus effects an increase in amplitude, the transients cause an additional delay, and when the stimulus

effects a decrease in amplitude, the transients cause an additional advance of the rhythm; only when the amplitude is not affected at all will there be an absence of transients in the resulting phase shift. As a result, the direction and amount of the phase shifts of the rhythm depend on the phase of the rhythm hit by the stimulus; different phase response curves result when the phase shifts are measured either immediately after the stimulus or in the steady state several cycles later.

The phase response curve, in total, depends primarily on the initial state of the rhythm. Figure 23 shows phase response curves (dotted lines, measured immediately following the stimulus; solid lines, measured in the steady state), measured with the same stimulus, in rhythms with three different constant levels of z (the left diagram originates from Fig. 22). The rhythm with the medium level, which has the largest amplitude and is least sensitive against perturbations (Fig. 9), shows the phase response curve with the smallest amplitude and the most sinusoidal shape. Moreover, the phase response curve is dependent on the parameters of the stimulus. Figure 24 shows phase response curves measured in the same rhythm (with $z = 2.1$) but with stimuli of varying duration (above) and intensity (below). In summary, the stronger the stimulus (in the duration and/or intensity), the larger is the amplitude of the generated phase response curve, and the more its shape deviates from the sinusoidal and tends to be sawtooth; with very strong stimuli, the range of the curve covers a full cycle, and at a certain phase it is ambiguous whether there is a phase delay or a phase advance.

From a phase response curve, some general statements can be made about the ability of the rhythm to become synchronized by corresponding periodic stimuli. For example, the amplitude of the phase response curve is positively

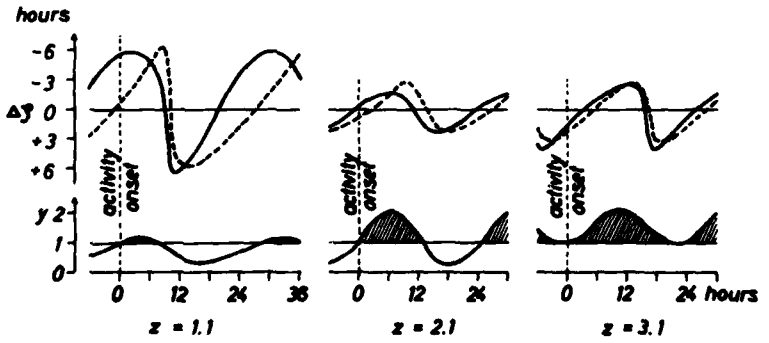


FIG. 23. Solutions of the model equation 7. Phase response curves (*upper diagrams*) computed according to Fig. 22, from oscillations with three different initial conditions (z values; lower diagrams, for comparison) but with identical stimuli. *Solid lines*, phase response curves measured many cycles after the perturbations (in the steady state). *Dotted lines*, phase response curves measured within the first cycle after the perturbations. (From Wever, ref. 25, with permission.)

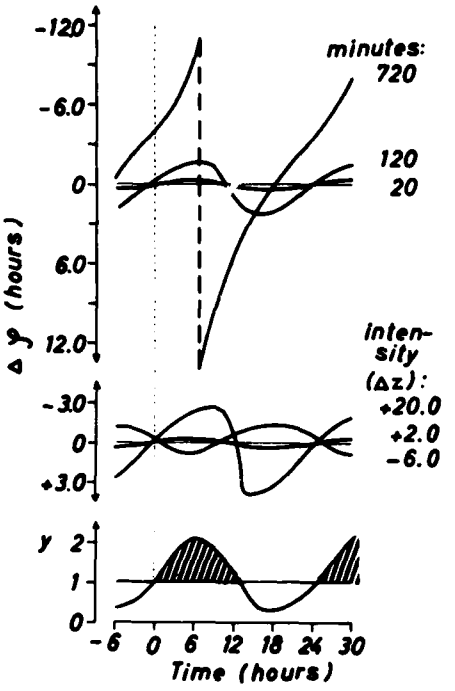


FIG. 24. Solutions of the model equation 7. Phase response curves, computed according to Fig. 22, from identical oscillations ($z = 2.1$) (lower diagram, for comparison) but with different stimuli. **Top:** "Light pulses" of different durations (always $\Delta z = 2$). **Middle:** Two "light pulses" of different intensity and one "dark pulse" (duration always 2 hr). (From Wever, ref. 25, with permission.)

correlated to the width of the range of entrainment (26). Or, the slope of the response curve at a certain phase indicates the mode of synchronization at this phase: with a slope less than 1, aperiodic adaptation to a steady state in the phase relationship is described; with a slope between 1 and 2, alternating but fading phase-angle differences are described; with a slope greater than 2, a steady state in the phase relationship cannot be realized (22).

These general statements are based on the fact that phase response curves and synchronization are both expressions of the phase-control ability of external forces. Special statements have also been made concerning phase response curves. The assumption is that from the phase response curve of an organism (as measured in the manner demonstrated in Fig. 22), one can compute phase-angle differences of circadian rhythms of the same organism to a stimulus given periodically. This assumption is based on two preconditions: (a) the phase response curve that is valid under the influence of a Zeitgeber is identical with that measured under constant conditions (as in Fig. 22), or can be derived by rules to be stated from the measured curve. (b) Preconditions concerning phase relationship with a special organism fit better when computed from a phase response curve of the same organism (or at least the same species) than when computed from any other phase response curve. The second precondition has been demonstrated only in special cases among very different

species. The first precondition can be shown to be doubtful on the basis of computations just discussed.

Phase response curves can be measured directly, not only under constant conditions but also under varying conditions (2). When doing so under the influence of a Zeitgeber, the most efficient method is to use a Zeitgeber outside the range of entrainment, hence producing relative coordination (*vide supra*). Figure 25 presents a population of computed phase response curves that are all based on the same rhythm (upper panel) and that are all computed with the same periodic alternation of the external force, but with slightly varying Zeitgeber periods. The solid lines represent computations with Zeitgeber periods outside the range of entrainment where relative coordination occurs; the dotted line comes from a Zeitgeber period inside the range of entrainment (computed during reentrainment after Zeitgeber shifts). It is obvious that all

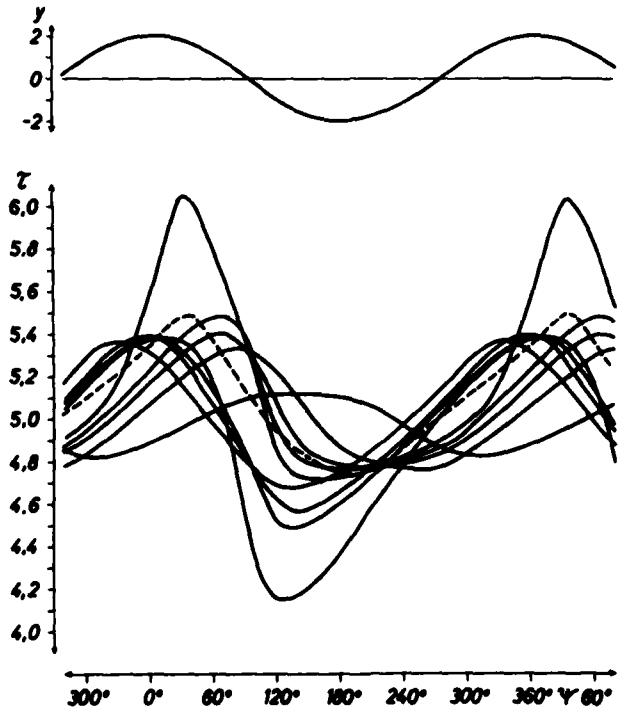


FIG. 25. Solutions of equation 2. Phase response curves from identical oscillations (*upper diagram*), computed under the influence of Zeitgebers with identical strengths but different periods [according to Aschoff (2), type IV]. *Solid lines*, Computed with Zeitgeber periods outside the range of entrainment (i.e., in the state of "relative coordination"). *Dotted line*, Computed with a Zeitgeber period in the middle of the range of entrainment after phase shifts of the Zeitgeber. (From Wever, ref. 39, with permission. Copyright: Academic Press Inc., London, Ltd.)

phase response curves differ considerably in amplitude, wave shape, and mean value; the differences between the individual curves are greater than differences in phase response curves when based on different rhythms (39). The width of the range of entrainment cannot be derived, as there is only one range of entrainment but many different ranges of phase response curves. The reason for the great differences within the population of phase response curves, measured with the same original rhythm and released by the same external stimulus, is the differential change in the rhythm's amplitude under the influence of these stimuli. All parameters of a phase response curve depend strongly on the amplitude of the original rhythm. And during the course of relative coordination, the amplitude of the original rhythm varies considerably (cf. Fig. 15) and very differently with different Zeitgeber periods.

SOLUTIONS OF EXTENDED MODEL EQUATIONS

Systems of Coupled Oscillators

A system consisting of two or more mutually coupled oscillators has a great degree of freedom and hence is capable of describing much more specific details of circadian rhythmicity than a single oscillator. Every additional degree of freedom reveals a great variety of possibilities in describing additional properties. Only to avoid ambiguities have the possibilities of a one-oscillator model been stressed in the preceding sections; this does not mean that multioscillator models should be neglected. Sufficient experimental evidence demands recognition of the multioscillator system (47).

Undoubtedly, every detail that can be described by one oscillator can be described as well by a two-oscillator or even a multioscillator system. The relevant question becomes how complicated details can also be described by a simple one-oscillator model having only one degree of freedom. Starting with the more complicated model raises the possibility of overlooking essential features of the simple model that, while being essential parts of the complicated model, cannot be studied reliably when only the complicated model is considered. It is therefore only after all possibilities of the one-oscillator model have been exhausted that, in a subsequent step, two or even more mutually coupled oscillators should be studied.

Out of the variety of solutions, which is much greater in the complex model than in the simple model, only some few examples, according to extensions of the simple model equation of the second order, will be presented. Much more than in previous discussions, the solutions cannot be complete; they can only suggest the directions in which the referred model extensions will lead. With this, of course, the rule is violated that the number of predictions deduced from the

solutions must be greater for several orders of magnitude than the number of preconditions put into the modeling process.

Third-Order Systems

In circadian rhythms of many organisms, a bimodal or even a multimodal wave shape is observed, rather than a monomodal one (4). Such shapes cannot, in principle, be described by an equation of the second order. Preceding the application of a multioscillator system, which is defined by the cooperation of at least two self-sustaining oscillators, the possibilities of the more simple extension to the third order should be evaluated. In Fig. 5, the most simple case of a self-sustaining oscillation of the third order (extended Van der Pol equation) has been shown. It shows the characteristics of the third-order supplementation; these characteristics maintain when the more appropriate equation 6 is supplemented by a corresponding term. It is characteristic that with a large enough ε value, the monomodal shape of the rhythm is exchanged for a multimodal shape. In general, the primary rhythm is superimposed by a secondary rhythm that has a higher frequency, where it is not capable of persisting self-sustainingly but damps out; the secondary rhythm starts oscillating only after it is pushed by the primary rhythm, and that is stronger the steeper the slope of the primary rhythm. In particular, this means in the case of $\varepsilon = 1$ (an appropriate approximation to biological results), that the shape is bimodal; with increasing external force (with shortening period), it alters from an "alternans" (secondary maximum preceding the primary) to a "bigeminus" (secondary maximum lagging behind the primary) (3).

Revised solutions to equation 8 are presented in Fig. 26 (slightly modified by a nonlinear restoring force) in the form of a multiple "actogram," as results from animal experiments are mostly presented; it shows the transient behavior of an activity–rest rhythm during a decrease of the external force (analogue to a decrease in light intensity). At day 10, the "activity episode" splits into two components. The stronger primary component holds its general pattern during the whole course; even before the split, the activity episode shows a bimodal pattern, with the higher peak at its beginning. The secondary component becomes smaller and smaller until it eventually disappears; it moves steadily away from its original primary component and gradually approaches the primary component of the following cycle. Consequently, the intervals between successive secondary components are, from its occurrence until its disappearance, consistently longer than the intervals between successive primary components because of the transiential change in the wave shape. With slight modifications of the coefficients, particularly an enlargement of ε, the originally uniform activity episode splits into three or more components (Fig. 5, with $\varepsilon = 10$); the general course of the then multiply split rhythm is similar to that shown in Fig. 26.

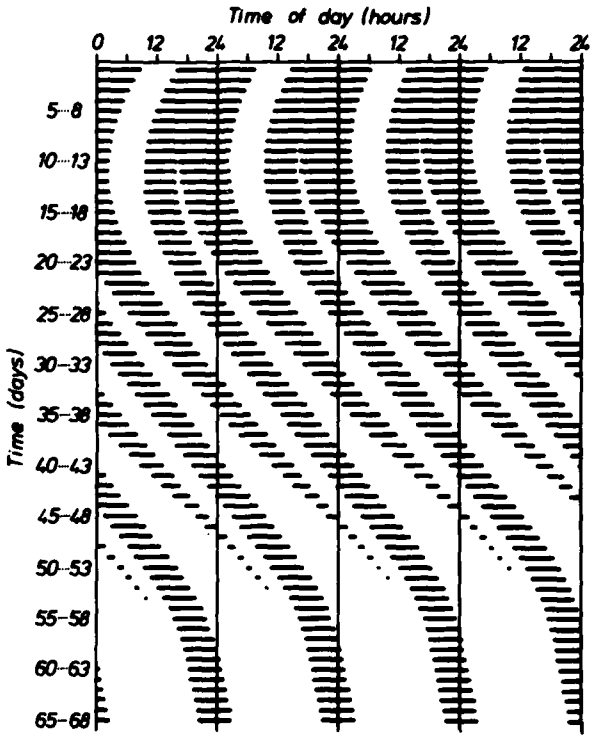


FIG. 26. Solutions of equation 8 (with $\varepsilon = 1$, and modified by slightly nonlinear restoring force) with a slight decrease in the external force. From the oscillation, only the "activity episodes" (above a threshold at $y = 0$) are drawn in the manner of biological data (quadruple plot).

Figure 26 resembles a phenomenon frequently observed in animal experiments, i.e., the "splitting" phenomenon. Either spontaneously or after changes in light intensity, a free-running activity–rest rhythm splits into two (or more, in rarer cases) activity components, with a simultaneous change in the overall period (10,11,13,14,16). After the splitting of the activity episode, the different components shift apart from each other, but never more than 180°. In the final steady state, therefore, the different components are synchronized with a constant mutual phase relationship. Consequently, these components show different intervals between successive cycles during the transient state (as they do in Fig. 26); if the precision of the rhythm is sufficiently high, the difference in intervals may be statistically significant. Correct period analyses do not result in separable periods; it would therefore be misleading to speak of different periods to be measured from the different components. The pitfall in the apparent estimation of a period is not readily obvious as it is in the comparable mathematical computation (Fig. 26). Its estimation would be based on

misapplication of statistics and on an inadequate definition of a period, similar to the error in period estimation mentioned in the context of Fig. 8. Extensive demonstration of both these cases should prevent incorrect conclusions being drawn from faulty analyses.

The diversity of possible solutions for model equations of the third order increases considerably with additional variations of the coefficient; then, the ratios in both period and amplitude of the primary and secondary components of the bimodal or multimodal rhythm vary (22). With further variations in other coefficients of equation 8, numerous phenomena can be described that are frequently coordinated to multioscillator systems. However, a third-order equation cannot generate a rhythm with two (or more) separate periods within the circadian range. Third-order systems should be classified, therefore, as one-oscillator systems.

Fourth-Order Systems

There have been experiments with humans clearly demonstrating the persistence of two or more steady-state oscillators with different periods, hence shifting apart from each other for multiples of 360° (47). To describe these results, both the simple equation 7 and the third-order equation 8 are insufficient; here, two (or more) separately self-sustaining oscillators have to be accepted that are mutually coupled; therefore, the system of equation 9 must be applied.

In order to reasonably restrict the number of free parameters in equation 9, ratios between the various coefficients will be determined from results of human experiments. With an iterative process, then, the application of these ratios will be tested with regard to their agreement with results of diverse experiments. Again, it is remarkable that consideration of but a few different complementary experimental results leads to the determination of such ratios. In this way, the establishment of an appropriate system of equations yields a reduction in the degrees of freedom.

In the case of spontaneous internal desynchronization, the periods of different overt rhythms shift in opposite directions, with a ratio of change always close to 1:12. Of course, the precondition for the occurrence of internal desynchronization is different intrinsic periods (or coefficients of frequency ω) of the two rhythms. The differential change in periods, however, can be described in alternative ways, with additional consequences. First, assuming corresponding different coupling coefficients c, the intrinsic amplitudes (or the coefficients a) will then be equal in both rhythms. Alternatively, equal coupling coefficients will lead to differing intrinsic amplitudes. With regard to the periods, both alternatives are compatible with biological results as well. Therefore, additional types of experiments must be considered in examining additional aspects of rhythms.

With a synchronizing external Zeitgeber, the external force either can affect

only one of the underlying oscillators (i.e., the "sleep–wake oscillator") or can affect both simultaneously. In the first case, the Zeitgeber will affect another oscillator (i.e., the "temperature oscillator") only indirectly, via the first oscillator. This assumption will necessarily lead to a correlation between the combined periods of the two rhythms and their mutual phase relationship, which is independent of the presence of a synchronizing Zeitgeber. In fact, human circadian rhythms consistently show a change in internal phase relationship when the period of a synchronizing Zeitgeber changes; on the other hand, a constant internal phase relationship is shown when the period changes under constant conditions (43). Therefore, these results do not agree with the necessary consequences of the first assumption. However, they are in agreement with the second assumption when the intrinsic amplitudes of the two rhythms are assumed to be different. Consequently, one can now discriminate between these alternatives concerning the opposite changes in period during spontaneous internal desynchronization.

Simultaneous consideration of both types of experiments leads to an unambiguous conclusion: In the two subequations of the system (equation 9), the right sides are equivalent; the environmental stimuli z, as well as the coupling coefficients c, are identical (or nearly identical). Apart from the frequency coefficients, the amplitude coefficients a must be larger in the "temperature oscillator" than in the "activity oscillator." Moreover, the coefficient of damping increment ε is set slightly larger in the "activity oscillator" ($\varepsilon \approx 2$) than the common value of $\varepsilon = 0.5$. The consequence of this selection of coefficient ratios is that they may be tested independently, under conditions where subjects are exposed to competing external stimuli of different modes, operating with different temporal schedules (49).

Figure 27 shows a solution of equation 9 where the two oscillations are presented separately. Because both oscillations are self-sustaining, the mutual interaction leads to mutual synchronization only within limited ranges of periods, similar to the effect of a Zeitgeber. In this special case, the two coefficients of frequency deviate from each other to such a degree that they are close to the mutual entrainment limit. In the upper diagram, both oscillations are synchronized. After 14 days, a small decrease in their amplitudes (a) leads to drastic alterations in periods, which no longer coincide (lower diagram). This decrease in amplitude and therefore in the mutual interaction is sufficient to deviate from the mutual range of entrainment. If the system were to be affected by random noise, there would be a great probability that the system would deviate from the mutual entrainment limit and become internally desynchronized. This behavior would be analogous to that of a single rhythm under the influence of a Zeitgeber with a period close to the entrainment limit (*vide supra*). Again, it would be unlikely that the internally desynchronized system would resynchronize spontaneously, as a consequence of another excessive random fluctuation, because of the difference in the ranges of holding and catching a rhythm (*vide supra*).

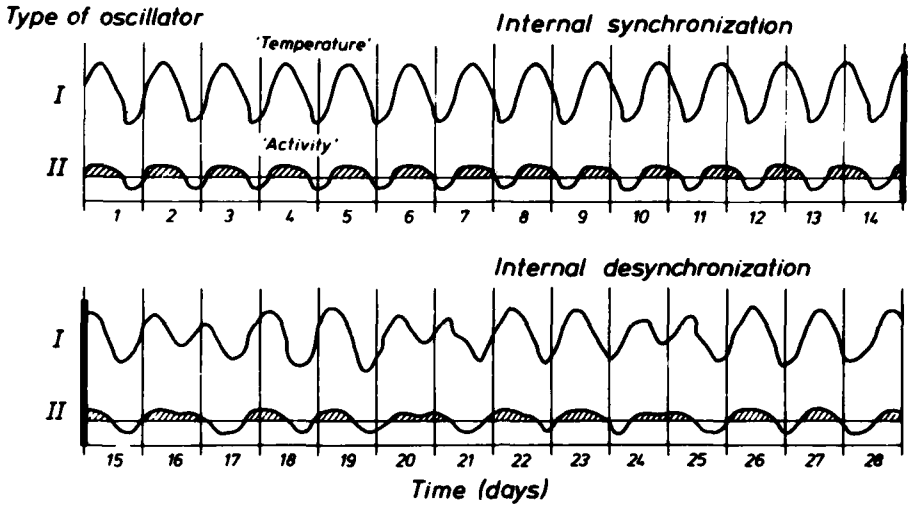


FIG. 27. Solutions of equation 9 computed with $\varepsilon_1 < \varepsilon_2$, $a_1 > a_2$, $\omega_1 > \omega_2$, and $c_1 = c_2$. The courses of the two suboscillations are drawn separately; the course of the type II oscillation is divided by a threshold in "activity" (*hatched areas*) and "rest." **Top:** both oscillations run in synchrony. **Bottom:** both *a* values are slightly decreased, and the oscillations do not run in synchrony.

In several respects, the separate rhythms in Fig. 27 behave as if under the influence of a Zeitgeber. In fact, there is no external Zeitgeber; however, each rhythm affects the other as an "internal Zeitgeber." Inside the mutual range of entrainment, the internal phase relationship between the two rhythms depends on the ratio between the intrinsic periods of the two rhythms (*vide supra*). Outside the mutual range of entrainment, the rhythms show the phenomenon of "internal relative coordination" (32) (*vide supra*), as expressed in the scalloping patterns of their phases; this is especially obvious in Fig. 28, where the same data are presented in a manner commonly used in human circadian rhythms. In special cases, even "internal relative entrainment" can be observed; this state presupposes sufficient differences in the amplitudes of the different rhythms. Moreover, as synchronization by an external Zeitgeber via phase-control is another expression for the existence of a "phase response curve" against the external stimulus under consideration (*vide supra*), mutual synchronization among different rhythms presupposes the existence of "mutual phase response curves." This means, the internal phase relationship between two rhythms (e.g., rhythms of deep body temperature and sleep–wake) determines the probability for the occurrence of special phases (e.g., minimum in temperature or sleep onset) and determines the duration of actual cycles and sections within a cycle. For instance, the interval between two successive minimum values in temperature (and the range of the temperature cycle)

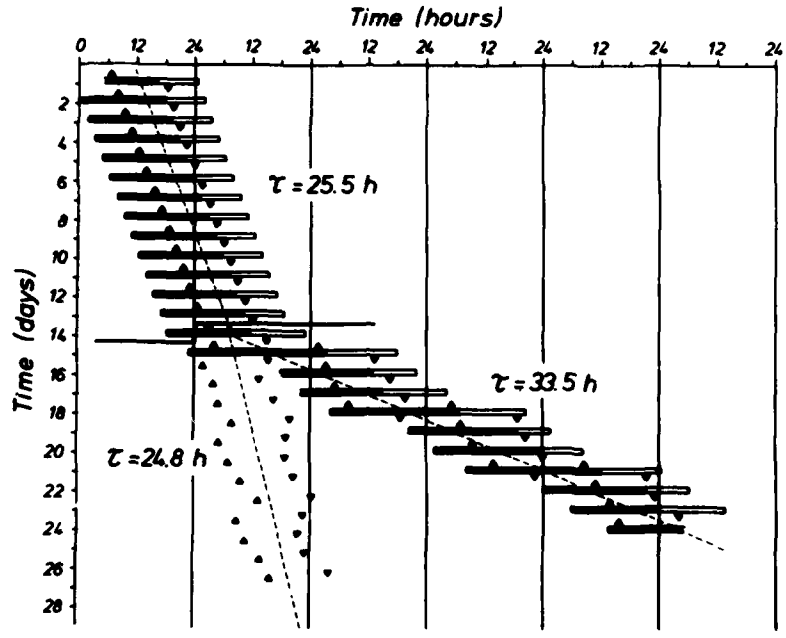


FIG. 28. Same solutions of equation 9 as in Fig. 27, but drawn in the manner of human data. The "activity rhythm" is represented by bars (*solid*, activity; *open*, rest), and the "rectal temperature rhythm" is represented by triangles indicating the temporal positions of the extremum values (*upright triangles*, maxima; *inverted triangles*, minima; *open triangles*, temporally correct redrawings of corresponding solid triangles).

depends on the actual position of the temperature cycle within the sleep–wake cycle; the duration of a sleep or wake episode depends systematically, i.e., in a predictable manner, on the actual position of sleep or wake onset relative to the minimum in deep body temperature. The corresponding correlations as measured in human circadian rhythms (60) are, therefore, consequences of the existence of mutual phase response curves and, hence, are necessary consequences of the ability of the rhythms to synchronize mutually.

In addition to the "oscillatory interaction" between the two rhythms, each rhythm affects the other directly via the "internal masking effect" (55) (*vide supra*). For instance, body temperature is generally higher during wakefulness than during sleep, independent of the mutual phases; this is particularly obvious when comparing corresponding extremum values of temperature during both states (Fig. 27). Also the internal masking effect contributes to the phenomenon of scalloping in the pattern of the phases.

Figures 27 and 28 demonstrate that this interaction leads to mutual synchronization only when the oscillators that are involved in the coupled system remain close in period (within a mutual range of entrainment). Because

these oscillators are nonlinear, there are also secondary ranges of entrainment (*vide supra*). Outside the primary and secondary ranges of entrainment, the mutual interaction does not lead to mutual synchronization. When the hypothesis is introduced that the separable oscillators originate from two different populations with different "oscillatory strengths" and, correspondingly, different standard deviations in their coefficients of frequency (47), the mutual interaction according to equation 9 leads to a concentration of the periods. Figure 29 demonstrates this effect. The dotted lines represent the hypothetical normal distributions in the periods of the oscillators of types I and II (nomenclature according to ref. 47) when the mutual interaction has ceased (i.e., with $c_1 = c_2 = 0$). After initiating the mutual coupling, the stronger oscillators of type I ("temperature oscillators") do not considerably alter the distribution of their periods. Conversely, the weaker oscillators of type II ("activity oscillators") alter the distribution of their periods considerably, as shown by the solid lines in Fig. 29. The arrows indicate the synchronizing influence of type I oscillators on type II oscillators within limited primary and secondary ranges of entrainment. In the distribution, the corresponding ranges of periods become empty, whereas other parts of the previous normal distribution are left. The result is a multimodal distribution of periods of type II

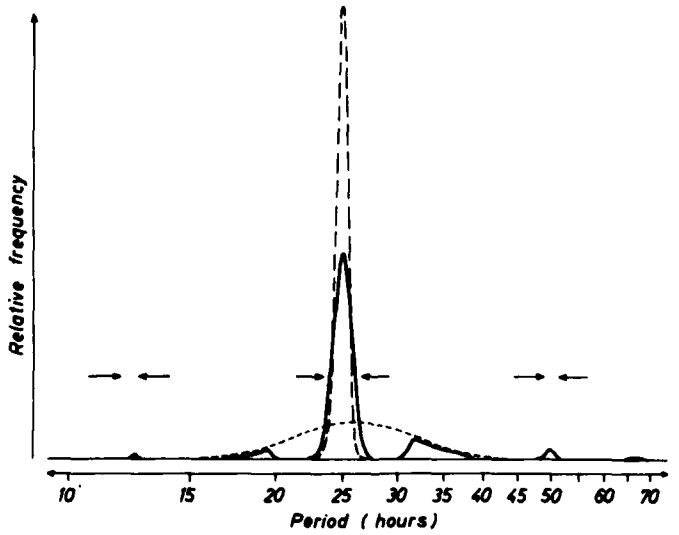


FIG. 29. Effect of the mutual interaction between the subsystems of equation 9 on the distribution of periods. *Dotted lines*, hypothetical distributions in periods of type I (narrow distribution) and type II (wide distribution) oscillations without mutual coupling. *Solid line*, distribution of periods in type II oscillations after introducing a mutual (symmetrical) coupling. *Arrows*, changes in the periods due to the interaction. (From Wever, ref. 47, with permission.)

oscillators. It is obvious that the multimodality of the resulting distribution is possible only because of their self-sustainment capacity, not only of the stronger type I oscillators but also of the weaker type II oscillators (47).

The self-sustainment capacity of the two oscillatory components is achieved in a manner similar to that for simple oscillators (*vide supra*), only within a limited oscillatory range. This range, however, is no longer defined by only that part of the external force that comes from the environment (z) but, in addition, by the influence of the other oscillator. In special solutions, only one oscillator is self-sustaining; it will drive the other rhythm, producing a mutual masking effect independent of its self-sustainment capacity.

As solutions of equation 9 are applicable to human circadian rhythms, they must be computed with the constant external force and with periodically alternating z. The total system of the two coupled oscillators (or oscillatory components) is then under the influence of an external Zeitgeber. Synchronizations of the two oscillators do not necessarily occur simultaneously; depending on the frequency of coefficients and the period of the Zeitgeber, only one of the two subsystems can be synchronized, while the other free-runs ("partial synchronization") (49). This state is mainly based on the differential widths of the ranges of entrainment of the two subsystems due to the differential oscillatory strengths.

Finally, the frequency coefficients in the two components of equation 9 must not be of the same order of magnitude. It is also possible to describe the mutual interaction of components with considerably differing periods. Here, the interaction between circadian and ultradian rhythms in humans (e.g., the 90-min sleep rhythm) is considered. The higher-frequency rhythm is then necessarily modulated in frequency and amplitude by the slower rhythm. Moreover, there exist solutions where the higher-frequency rhythm oscillates only at certain phases of the slower rhythm, e.g., during rest time, or during phases with lower (or higher) temperatures; it damps out at other parts of the cycle.

In summary, the number of possibilities with the multioscillator system (equation 9) is a high multiple of that with the one-oscillator system (equation 7). Additional solutions of the larger system that were necessary in order to reach a similar degree of completeness as in case of the single system would be far beyond the scope of this chapter. Moreover, they would not be meaningful, because they would lead to a host of predictions even for relatively simple problems, whereas the one-oscillator model (equation 7) offers one testable prediction for every problem.

Last, but not least, a very different aspect of the multioscillator system should be mentioned. A system of two mutually coupled oscillators according to equation 9 had been shown earlier to be of great interest in solving problems of homeostasis (23,29). The same terms that assist in stabilizing the period of the generated rhythms (*vide supra*) likewise guarantee stabilization of the mean value; in particular, this is the nonlinearity in the restoring force. Considering

this equivalence, two main (and apparently contradictory) principles in the dynamics of biological systems, rhythmicity and homeostasis, appear to melt into one another. They seem to be rather two complementary aspects of one general biological principle.

High-Frequency Rhythms

The purpose of this chapter is to establish a mathematical model for circadian rhythmicity; the deduced model equations, in fact, have the capacity to describe the dynamics of this rhythmicity. However, of additional interest is to examine the same model in its efficiency to describe different biological rhythmicities as well. In this regard, the applicability of the model will be extended, and its soundness can be confirmed as well. In order to examine this possibility, an extension of the original model equation had been offered that includes a coefficient of frequency (*vide supra*). In the following, solutions of equation 10 with high frequencies are presented that may describe phenomena of the central nervous system (24).

At first, solutions of equation 10 will be presented with $\omega = 10^8$. Solutions inside and outside of the oscillatory range are of equal interest; the limits of this range are identical with those of equation 7, i.e., the lower limit is at $z = 0.847$. Because of the high value of the effective coefficient of damping increment ($\varepsilon = 100$), the lower limit of the range of periodic adaptation is nearly identical ($z = 0.846$). Figure 30 shows solutions inside the oscillatory range with four different z values; it shows a series of spikes, the shape of which is independent of z, and the frequency or rate of which varies considerably with slightly varying z. Figure 31 shows solutions outside the oscillatory range (and the range of periodic adaptation) with $z = 0.65$, under the influence of different stimuli (short increases of z; the length of the arrow indicates the strength of the stimulus). The upper row shows the effects of a single stimulus: If a stimulus is

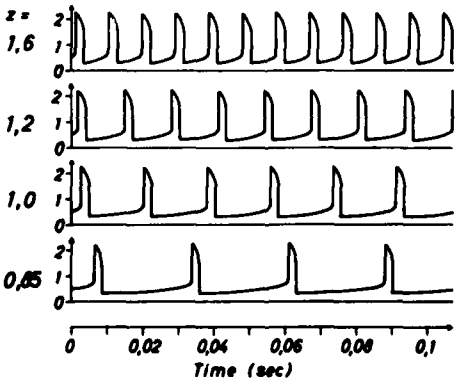


FIG. 30. Solutions of equation 10 with $\omega = 10^8$ and four different z values inside of the oscillatory range. (From Wever, ref. 24, with permission.)

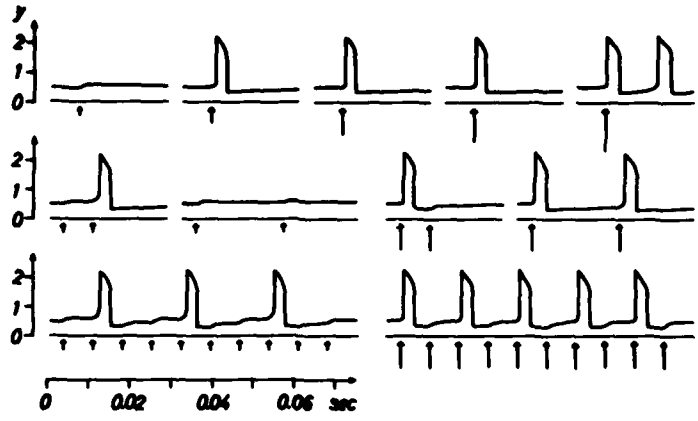


FIG. 31. Solutions of equation 10 with $\omega = 10^8$ and a z value outside the oscillatory range ($z = 0.65$) but under the influence of different stimuli (short-term increases of z). *Upper row*, influence of the strength of solitary stimuli (the length of the *arrow* indicates the increase in z). *Middle row*, influence of combined stimuli of different strengths. *Lower row*, influence of periodically repeated stimuli of different strengths. (From Wever, ref. 24, with permission.)

strong enough, it releases a reaction that is then independent of the stimulus. The middle row shows effects of two combined stimuli, the left part corresponding to the summation effect, and the right demonstrating a refractory phase. The lower row shows effects of periodically repeated stimuli: When the stimuli are too frequent to affect one reaction per stimulus, there is frequency demultiplication. In summary, equation 10, with $\omega = 10^8$, seems to be an appropriate model for describing the activity of nerve cells, either spontaneous (Fig. 30) or reactive (Fig. 31).

Reactive and spontaneous actions can merge into each other when the external force z is still closer to the limit of the range of entrainment. Figure 32 shows solutions of equation 10 with $z = 0.848$ (this time, ω is slightly smaller). The different lines show the effects of different stimuli (arrows) where the first stimulus remains the same. The comparison of lines (a) and (b) once more demonstrates the existence of a refractory phase. In line (c), a weaker stimulus at the same position as the second stimulus in (b) does not produce a reaction; however, it does release a reaction when occurring still later, as demonstrated in line (d). When no second stimulus follows at all (e), eventually a reaction occurs spontaneously (idle action). In lines (a) and (c), the dotted lines indicate the occurrence of the next reaction that was not released, in these cases, by a stimulus. They demonstrate that the second stimulus, though not releasing a direct reaction, was not completely ineffective; it releases a spontaneous reaction that is earlier than it would be without the stimulus [compare line (e)]. In summary, the readiness for the release of a reaction by a stimulus increases

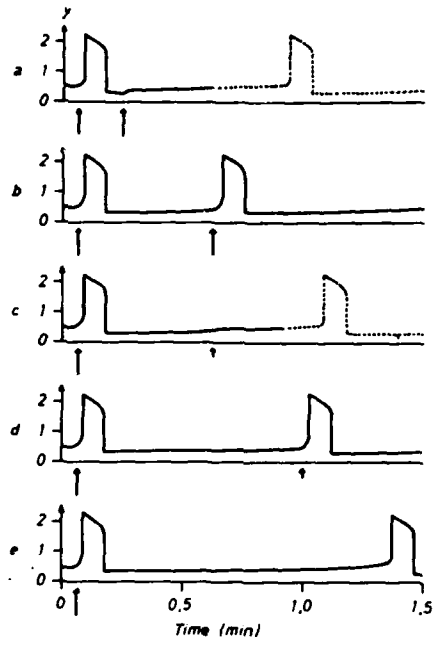


FIG. 32. Solutions of equation 10 with $\omega = 10^6$, a z value just inside of the oscillatory range ($z = 0.848$), and the additional influence of combined stimuli. After an initial stimulus (equal strength in every row), a second stimulus of varying strength and with varying interval is given. The dotted lines indicate the "spontaneous" occurrence of the next "reaction" after a "subliminal stimulus." (From Wever, ref. 24, with permission.)

with increasing interval following the preceding reaction, or the threshold for the release of the reaction lowers, eventually down to zero. These solutions may be adequate models for the release of instinctive actions (24).

CONCLUSIONS

In summarizing the results of the preceding sections, a simple oscillation equation that is based on very few preconditions is capable of predicting numerous features of circadian rhythmicity under a great variety of external conditions. The model equation has been deduced using alternative hypotheses tested in biological experiments. It is therefore meaningful to ask only to what degree it is sufficient, rather than if it is right or wrong. The equation in its simple form contains only one free parameter when all other coefficient values are fixed according to results of relevant experiments.

A sound agreement between model predictions and experimentally derived properties of circadian rhythms indicates that simple laws of oscillation theory govern the apparently very complex rhythmicity in the behavior of living organisms, including humans. It is tempting to conclude from this result that the basic structural mechanisms of generating and controlling circadian rhythms are also simple; and the evaluated dynamics of biological systems should assist in discovering these mechanisms. Therefore, it should not be hopeless to transpose

the mathematical parameters to structural properties. An example of such a correlation may be the coordination of the "oscillatory strength," or the rhythm's amplitude, to the number of cells cooperating in the establishment of a "pacemaker" in the circadian system (49). Following Wiener (57), the larger the number of mutually synchronizing "oscillatory cells," the more precise is the resulting "oscillatory center"; this principle has been introduced in circadian research by Barlow (7). In the human multioscillator system this principle means that the differential oscillatory strengths of different types of oscillators are due to the differential numbers of oscillatory cells cooperating in the constitution of an oscillatory center (49).

The primary basis for the modeling considerations as discussed in this chapter is the one-oscillator system. It is only such a simple system, with not more than one degree of freedom, that leads to unambiguous, testable predictions. Evidently, results from human experiments necessitate the concept of two-oscillator or even multioscillator models. The increasing degree of freedom in such extended systems can only lead to ambiguous predictions; with small variations in free parameters, a great variety of rhythm patterns can be verified. Within a multioscillator system, the properties of the single oscillators determine the behavior of the combined system. Therefore, thorough knowledge of all properties of the constituent single oscillators is a necessary precondition for the understanding of a multioscillator system. In addition, the interactions between the single oscillators determine the behavior of the combined system. This is why the degrees of freedom increase considerably when the increase in the number of participating oscillators is slight.

If predictions should be made concerning a two-oscillator system, the degrees of freedom must be considerably reduced. In the single-model equation, the coefficients of the various terms were set according to results of many diverse biological experiments, so that only the external force was a free parameter. Correspondingly, it has been shown in human experiments that in this way the determination of ratios between corresponding coefficients in the two equations is possible. Here, the coefficients in an equation that corresponds to "type I oscillator" rhythms can be set according to the coefficients in the single-model equation. Continuing in this manner, the use of results from more sophisticated experiments will eventually lead to unambiguous predictions from the two-oscillator model. Because in humans behavioral aspects control to a significant degree circadian rhythmicity (37,48), their consideration is of special relevance when the multioscillatory character of the circadian system is examined (54).

The model equation with all its extensions, as discussed in this chapter, was established in a series of papers about 20 years ago; simultaneously, the mathematical bases have been developed (20–29). The system of two coupled oscillators had been introduced originally to describe complicated (multimodal) wave shapes (21) and special problems of homeostasis (23,29). The model solutions presented here originate in large part from that time; the original solutions concerned the behavior of rhythms in the steady state. Several aspects of the model solutions were discovered only several years later (38–47),

including the application of the two-oscillator model to internally desynchronized rhythms; and interesting new properties of generated model rhythms continue to be developed (52–56). Hence, the process of evaluating solutions under continuously varying conditions is still in progress. It facilitates the selection of preferences for time-consuming experimental research, and it assists in formulating new hypotheses about the substantial basis of circadian rhythmicity.

It is remarkable that until the present time there has been no need to modify the original model equation, in spite of the fact that new experimental approaches have continually been applied to test the model predictions. Nevertheless, the model equation must be subject to further specifications. For example, when squares of the variable y are given, the only meaning is that the term under discussion has to increase more than proportional to y; differing powers, or even other functions like an exponential (with, of course, correspondingly differing numerical coefficients), may fit the experimental results as well. More important, the appropriateness of the mathematical model presented in this chapter must not preclude its being subjected to rigorous criticism. It is the aim of this chapter to demonstrate that this kind of successive trial-and-error modification will lead to a successful model equation; the same method when starting from a fundamentally differing base may lead to a very different but likewise sufficient model. The modeling process, in general, however, can never be assumed to be completed, because it is dynamic in itself.

REFERENCES

1. Aschoff, J. (1960): Exogenous and endogenous components in circadian rhythms. *Cold Spring Harbor Symp. Quant. Biol.*, 25:11–27.
2. Aschoff, J. (1965): Response curves in circadian rhythmicity. In: *Circadian Clocks*, edited by J. Aschoff, pp. 95–111. North Holland, Amsterdam.
3. Aschoff, J. (1966): Circadian activity pattern with two peaks. *Ecology*, 47:657–662.
4. Aschoff, J., and Honma, K. (1959): Art- und Individualmuster der Tagesperiodik. *Z. Vergl. Physiol.*, 42:383–392.
5. Aschoff, J., Gerecke, U., Kureck, A., Pohl, H., Rieger, P., von Saint Paul, U., and Wever, R. (1971): Interdependent parameters of circadian rhythms in birds and man. In: *Biochronometry*, edited by M. Menaker, pp. 3–27. National Academy of Sciences, Washington, D.C.
6. Aschoff, J., and von Saint Paul, U. (1976): Brain temperature in the unanaesthetized chicken: Its circadian rhythm of responsiveness to light. *Brain Res.*, 101:1–9.
7. Barlow, J. S. (1960): Discussion remark. *Cold Spring Harbor Symp. Quant. Biol.*, 25:54–55.
8. Brown, F. A., Jr. (1965): A unified theory for biological rhythms. In: *Circadian Clocks*, edited by J. Aschoff, pp. 231–261. North Holland, Amsterdam.
9. Brown, F. A., Jr. (1972): The 'clock' timing biological rhythms. *Am. Sci.*, 60:756–766.
10. Hoffmann, K. (1970): Circadiane Periodik bei Tupajas (Tupaia glis) in konstanten Bedingungen. *Zool. Anz. [Suppl.]*, 33:171–177.
11. Hoffmann, K. (1971): Splitting of the circadian rhythm as a function of light intensity. In: *Biochronometry*, edited by M. Menaker, pp. 134–146. National Academy of Sciences, Washington, D.C.
12. von Holst, E. (1939): Die relative Koordination als Phaenomen und als Methode zentralnervoeser Funktionsanalyse. *Ergeb. Physiol.*, 42:228–306.

13. Pittendrigh, C. S. (1960): Circadian rhythms and the circadian organization of living systems. *Cold Spring Harbor Symp. Quant. Biol.*, 25:159–184.
14. Pittendrigh, C. S. (1967): Circadian rhythms, space research and manned space flight. *Life Sci. Space Res.*, 5:122–134.
15. Pittendrigh, C. S., and Daan, S. (1976): A functional analysis of circadian pacemakers in nocturnal rodents. I. The stability and lability of spontaneous frequency. *J. Comp. Physiol. A*, 106:223–252.
16. Pittendrigh, C. S., and Daan, S. (1976): A functional analysis of circadian pacemakers in nocturnal rodents. V. Pacemaker structure: A clock for all seasons. *J. Comp. Physiol. A*, 106:333–355.
17. Van der Pol, B. (1926): Ueber "Relaxationsschwingungen." *Jb. Drahtl. Telegr. Teleph.*, 28:178–184.
18. Van der Pol, B. (1927): Ueber "Relaxationsschwingungen" II. *Jb. Drahtl. Telegr. Teleph.*, 29:114–118.
19. Van der Pol, B. (1934): The nonlinear theory of electric oscillations. *Proc. I.R.E.*, 22:1054–1086.
20. Wever, R. (1960): Possibilities of phase-control, demonstrated by an electronic model. *Cold Spring Harbor Symp. Quant. Biol.*, 25:197–201.
21. Wever, R. (1962): Zum Mechanismus der biologischen 24-Stunden-Periodik. *Kybernetik*, 1:139–154.
22. Wever, R. (1963): Zum Mechanismus der biologischen 24-Stunden-Periodik. II. Mitteilung. Der Einfluss des Gleichwertes auf die Eigenschaften selbsterregter Schwingungen. *Kybernetik*, 1:213–231.
23. Wever, R. (1963): Zum Problem der Regelung in der Biologie. *Pfluegers Arch.*, 278:89–90.
24. Wever, R. (1964): Ein mathematisches Modell fuer biologische Schwingungen. *Z. Tierpsych.*, 21:359–372.
25. Wever, R. (1964): Zum Mechanismus der biologischen 24-Stunden-Periodik. III. Mitteilung. Anwendung der Modell-Gleichung. *Kybernetik*, 2:127–144.
26. Wever, R. (1965): A mathematical model of circadian rhythms. In: *Circadian Clocks*, edited by J. Aschoff, pp. 47–63. North Holland, Amsterdam.
27. Wever, R. (1965): Pendulum versus relaxation oscillation. In: *Circadian Clocks*, edited by J. Aschoff, pp. 74–84. North Holland, Amsterdam.
28. Wever, R. (1966): The duration of re-entrainment of circadian rhythms after phase shifts of the zeitgeber. A theoretical investigation. *J. Theor. Biol.*, 13:187–201.
29. Wever, R. (1966): Ein mathematisches Modell fuer die circadiane Periodik. *Z. Angew. Math. Mech.*, 46:148–157.
30. Wever, R. (1967): Zum Einfluss der Daemmerung auf die circadiane Periodik. *Z. Vergl. Physiol.*, 55:255–277.
31. Wever, R. (1967): Modelle circadianer Rhythmen. *Nachr. Akad. Wiss. Goettingen*, 10:129–131.
32. Wever, R. (1968): Einfluss schwacher elektro-magnetischer Felder auf die circadiane Periodik des Menschen. *Naturwissenschaften*, 55:29–32.
33. Wever, R. (1968): Mathematical models of circadian rhythms and their applicability to men. In: *Cycles biologiques et psychiatrie*, edited by J. de Ajuriaguerra, pp. 61–72. Masson & Cie, Paris.
34. Wever, R. (1968): Gesetzmaessigkeiten der circadianen Periodik des Menschen, geprueft an der Wirkung eines schwachen elektrischen Wechselfeldes. *Pfluegers Arch.*, 302:97–122.
35. Wever, R. (1969): Autonome circadiane Periodik des Menschen unter dem Einfluss verschiedener Beleuchtungs-Bedingungen. *Pfluegers Arch.*, 306:71–91.
36. Wever, R. (1969): Untersuchungen zur circadianen Periodik des Menschen mit besonderer Beruecksichtigung des Einflusses schwacher elektrischer Wechselfelder. Bundesminist. wiss. Forschg., Forschungsber. W 69–31.
37. Wever, R. (1970): Zur Zeitgeber-Staerke eines Licht-Dunkel-Wechsels fuer die circadiane Periodik des Menschen. *Pfluegers Arch.*, 321:133–142.
38. Wever, R. (1971): Influence of electric fields on some parameters of circadian rhythms in man. In: *Biochronometry*, edited by M. Menaker, pp. 117–132. National Academy of Sciences, Washington, D.C.

39. Wever, R. (1972): Virtual synchronization towards the limits of the range of entrainment. *J. Theor. Biol.*, 36:119–132.
40. Wever, R. (1972): Mutual relations between different physiological functions in circadian rhythms in man. *J. Interdiscipl. Cycle Res.*, 3:253–265.
41. Wever, R. (1972): Ein mathematisches Modell fuer die biologische Tagesperiodik. In: *Akustik und Schwingungslehre*, pp. 512–515. VDE-Verlag, Berlin.
42. Wever, R. (1973): Reactions of model-oscillations to external stimuli depending on type of oscillation. In: *Biogenesis-Evolution-Homeostatis*, edited by A. Locker, pp. 81–83. Springer-Verlag, Berlin.
43. Wever, R. (1973): Internal phase-angle differences in human circadian rhythms: Causes for changes and problems of determinations. *Intern. J. Chronobiol.*, 1:371–390.
44. Wever, R. (1973): Der Einfluss des Lichtes auf die circadiane Periodik des Menschen. I. Einfluss auf die autonome Periode. *Z. Physik. Med.*, 3:121–134.
45. Wever, R. (1974): Die biologische Zeituhr. *Koelner Techn. Mitteil.*, 89/6:1–4.
46. Wever, R. (1974): Different aspects of the studies of human circadian rhythms under the influence of weak electric fields. In: *Chronobiology*, edited by L. E. Scheving, F. Halberg, and J. E. Pauly, pp. 694–699. Igaku Shoin, Tokyo.
47. Wever, R. (1975): The circadian multi-oscillator system of man. *Int. J. Chronobiol.*, 3:19–55.
48. Wever, R. (1978): Grundlagen der Tagesperiodik beim Menschen. In: *Rhythmusprobleme in der Psychiatrie*, edited by H. Heimann and B. Pflug, pp. 1–23. Gustav Fischer Verlag, Stuttgart.
49. Wever, R. A. (1979): *The Circadian System of Man*. Springer-Verlag, Berlin.
50. Wever, R. A. (1980): Circadian rhythms of finches under bright light: Is self-sustainment a precondition for circadian rhythmicity? *J. Comp. Physiol.*, A139:49–58.
51. Wever, R. A. (1980): Circadian rhythms of finches under steadily changing light intensity: Are self-sustaining circadian rhythms self-excitatory? *J. Comp. Physiol.*, A140:113–119.
52. Wever, R. (1980): Die Tagesperiodik des Menschen—Grundlagen und Probleme. *Betriebsaerztliches*, 80/1:1–33.
53. Wever, R. A. (1980): Phase shifts of human circadian rhythms due to shifts of artificial zeitgebers. *Chronobiologia*, 7:303–327.
54. Wever, R. A. (1982): Behavioral aspects of circadian rhythmicity. In: *Rhythmic Aspects of Behavior*, edited by F. M. Brown and R. C. Graeber, pp. 105–171. L. Erlbaum Associates, Hillsdale, N.J.
55. Wever, R. A. (1983): Organization of the human circadian systems: Internal interactions. In: *Circadian Rhythms in Psychiatry*, edited by T. A. Wehr and F. K. Goodwin, Boxwood Press, Neuroscience Series, Los Angeles.
56. Wever, R. A. (1983): The problem of long-term stability in circadian rhythms (*in preparation*).
57. Wiener, N. (1958): *Nonlinear Problems in Random Theory*. Wiley, New York.
58. Zulley, J. (1976): Schlaf und Temperatur unter freilaufenden Bedingungen. *Ber. 30. Kongr. Dtsch. Ges. Psychol.*, pp. 398–399.
59. Zulley, J. (1979): *Der Einfluss von Zeitgebern auf den Schlaf des Menschen*. Fischer-Verlag, Frankfurt am Main.
60. Zulley, J., Wever, R., and Aschoff, J. (1981): The dependence of onset and duration of sleep on the circadian rhythm of rectal temperature. *Pfluegers Arch.*, 391:314–318.

DISCUSSION

Dr. Kronauer: I would like to make one comment and ask a question. The comment is in defense of the mathematician. The phenomenon that you describe as relative entrainment was actually described in Stokler's very excellent book on nonlinear oscillators in 1949, and, for all I know, it may have been described earlier. There is one question I would like to ask you that has always puzzled me, because you use as the excitation on the right-hand side of your equation a combination of z, its first derivative,

and its second derivative. If you put a square wave in, of course, that gives you a delta function and a second-order delta function at every place the square wave comes. And so, first of all, do you really think that there is a mechanism that converts steps of light into impulses, and, second, how do you handle it mathematically in your differential equation routine?

Dr. Wever: The answer is simple. I cannot deal with step functions. But in nature we never have steps, so we have only more or less slow increases, and then I can calculate the duration of a twilight; and in fact from biological experiments we have clear indications that the twilight transition has a remarkably large influence on many different parameters of the entrained rhythm. You are right that I cannot compute a rectangular cycle, but I don't see any reason for trying to do it.

Dr. Edmunds: It may be heresy, but I do not think unicells ever sleep or wake or either do all one or all the other. In your models dealing with the activity–rest ratio, would you eliminate the threshold in order to model unicells?

Dr. Wever: The threshold had been introduced secondarily to describe additional features like sleep–wake alternation; it is not a constituent of the original model. I have the feeling that introduction of a threshold separating these two states is a very late step in evolution. I do not need the threshold for 90% of what I am saying, but for the last 10% dealing with the sleep–wake cycle, I need such a threshold.

Dr. Moore-Ede: I would like to have you clarify why you see it as so unacceptable to have a function which goes alternately negative and positive as an oscillation. You made the point that you do not see negative values in body temperature. I agree with you, but on the other hand, in neural systems we clearly see changes in membrane potential which can be either positive or negative during the course of an action potential. So I see no fundamental reason why you cannot have a function that moves either side of a zero line.

Dr. Wever: We have secured entrainment because of the second nonlinearity. As a result, we have simultaneously parametric as well as nonparametric entrainment. When you look solely for the loss of parametric entrainment, you can see that you have a bistable phase when the rhythm is alternating between positive and negative values, but you have an unambiguous phase relation when you have a consistent sign. That is one of the reasons. All these modifications assist in stabilizing the generated rhythm in spite of the superimposed value fluctuations.

Dr. Weitzman: Along those lines, just for clarification, you know that some physiological functions do reach zero values for periods of time. For example, cortisol is totally not secreted for many hours and therefore remains at zero, even though it never goes negative, obviously. How do you deal with that?

Dr. Wever: I cannot give you an answer. Please consider that my equations are more than 20 years old. At that time we knew nothing about cortisol, so we were surprised that most of the data fit. Maybe it would be a good idea to modify the equations, but I don't think that is a very important point in the general picture. This equation should show nothing but the very general behavior—nothing more, not specifically cortisol or anything else.

Dr. Moore-Ede: You made the comment that you do not like coupled oscillators, if I understand what you said. Could you explain that?

Dr. Wever: What I do not like is to apply coupled oscillators before the potentialities of single oscillators have been exhausted. That is, because they have such a tremendous number of free parameters. What I would like is a model with only one free parameter. Then we get an unambiguous answer. A system of two coupled oscillators has six free parameters. The combination of these six parameters gives us such a tremendous number

of answers; so when we get a fit between a model simulation and a biological result, it is hard to decide whether or not it is only due to a good choice of the parameters with a very poor equation.

Dr. Moore-Ede: You are saying that you view the human system as a multioscillator system, but at the same time, you are not modeling a multioscillator system?

Dr. Wever: Of course, results of human experiments force the assumption of a multioscillator system, and my multioscillator model fits the data very well. However, the consideration of only coupling properties says nearly nothing about the structure of each single oscillator. It says, in this respect, much more when parameters of a one-oscillator model are estimated, as the basis also of combinations of oscillators.

Dr. Edmunds: You do not like multioscillator systems because they are hard to model, but in reality the experimental data clearly show coupling interactions. Therefore, it is a problem in trying to model these things. Just because you cannot model them does not mean that they are not good.

Dr. Wever: Multioscillator systems are not hard to model, but agreement between model simulations and biological results is hard to interpret in multioscillator systems. Rather, multioscillator systems are easy to model, because agreement with every experimental result can be achieved with various combinations of the many parameters. My only point is that we need a great amount of biological data to differentiate between all these possibilities; and in several respects, we have not yet enough of them.

Mathematical Models of the Circadian Sleep-Wake Cycle, edited by M. C. Moore-Ede and C. A. Czeisler. Raven Press, New York © 1984.

Are Separate Temperature and Activity Oscillators Necessary to Explain the Phenomena of Human Circadian Rhythms?

Charmane Eastman

University of Chicago Sleep Laboratory, Chicago, Illinois 60637

TWO TYPES OF MULTIPLE-OSCILLATOR THEORIES

It is generally accepted that circadian systems are composed of multiple oscillators. This chapter describes a single-oscillator model of human circadian rhythms. This does not necessarily represent a contradiction. The potential for confusion stems from the fact that there are two different types of multiple-oscillator theories based on different types of experimental evidence.

In the first type of theory, multiple oscillators interact to control a single rhythmic variable, usually locomotor activity (17,18,23,24,28,35–38,42, 50,51). Evidence for this type of organization is provided by various sets of data, including the "splitting" of locomotor-activity rhythms in various animal species (37), the breakdown of circadian rhythms after lesions of the suprachiasmatic nuclei (SCN) (38,42), coexisting free-running and food-entrained activity rhythms in rodents (12), and the demonstration of circadian rhythms in isolated mammalian organs and tissues (13,26,39,43).

In the second type of multiple-oscillator theory, separate oscillators are responsible for the control of different physiological and behavioral circadian rhythms. The strongest evidence for this type of multiple-oscillator organization has been "spontaneous internal desynchronization" in the human, which has been attributed to the uncoupling of separate temperature and activity oscillators (9,10,30,34,48,49) (see R. E. Kronauer, *this volume*).

The single-oscillator model presented here is an alternative to the second type of multiple-oscillator theory, in that all the various circadian rhythms are controlled by the same circadian oscillator. This model does not dispute the multiple-oscillator organization proposed by the first type of multiple-oscillator theory; in fact, the single oscillator may itself be composed of multiple oscillators (23,24). However, the model demonstrates that spontaneous internal desynchronization, as well as other human circadian phenomena, can be explained with the single oscillator remaining intact, i.e., without a desynchronization or breakdown of the circadian system into its component parts.

SPONTANEOUS INTERNAL DESYNCHRONIZATION

Humans in temporal isolation, like other animals, usually show synchronized, free-running circadian rhythms in all the various physiological and behavioral measures (9,10,34,49). An example is shown in Fig. 1, section A. For the first 14 days, the circadian rhythms of temperature and activity were internally synchronized; they free-ran with the same period (25.7 hr) and maintained fairly fixed internal phase relationships. Then, at about day 15, spontaneously, for no apparent reason, the rhythmic patterns changed dramatically. The periodograms in section B show that two periods were present in each rhythm, one at 33.4 hr that was dominant in the activity rhythm and one at 25.1 hr that was dominant in the temperature rhythm. The daily chart (Fig. 1, top) shows activity plotted to emphasize the 33.4-hr period and temperature plotted on the left (solid triangles) to emphasize the 25.1-hr period. Over the course of days, the activity and temperature rhythms assumed radically different internal phase relationships.

For many years, patterns like these have been known as spontaneous internal desynchronization, from the extensive work of Aschoff's group (9,10,48,49). Other investigators have found similar aberrations in the activity rhythm (other variables were not always measured), but they have not always labeled these patterns as spontaneous internal desynchronization (15,16,27,32,44,45). Aschoff's group has found that only about one-third of subjects become internally desynchronized. These investigators isolated subjects in constant conditions for about a month (9,10,48,49). Czeisler claims that when subjects are isolated for longer periods of time, at least 2 months, they all eventually desynchronize (16,30,45).

A summary of all the major periods revealed by periodograms from over 100 subjects isolated from time cues by Aschoff's group is shown in Fig. 2. During synchronized free runs (Fig. 2, top) the average period was about 25 hr. For each subject, both temperature and activity free-ran with the same period. During "true" spontaneous internal desynchronization (Fig. 2, bottom) the periodograms usually revealed two periods in each rhythm (Fig. 1). One was about 25 hr, falling into distribution C. This period was usually dominant in the temperature rhythm. The other was either very long, between about 30 and 40 hr, falling into distribution D, or very short, between about 15 and 20 hr, falling into distribution B. This period was usually dominant in the activity rhythm. Twenty-four percent of the subjects showed "true" internal desynchronization in which the two periods were not multiples of each other. Another 8% showed "apparent desynchronization," in which a bicircadian period (about 50 hr) or a semicircadian period (about 12.5 hr) appeared with the circadian period (of about 25 hr) (10,49).

The type of internal desynchronization described here has been seen only in humans. Moore-Ede's group has reported spontaneous internal desynchronization in squirrel monkeys maintained in constant light. However, in their

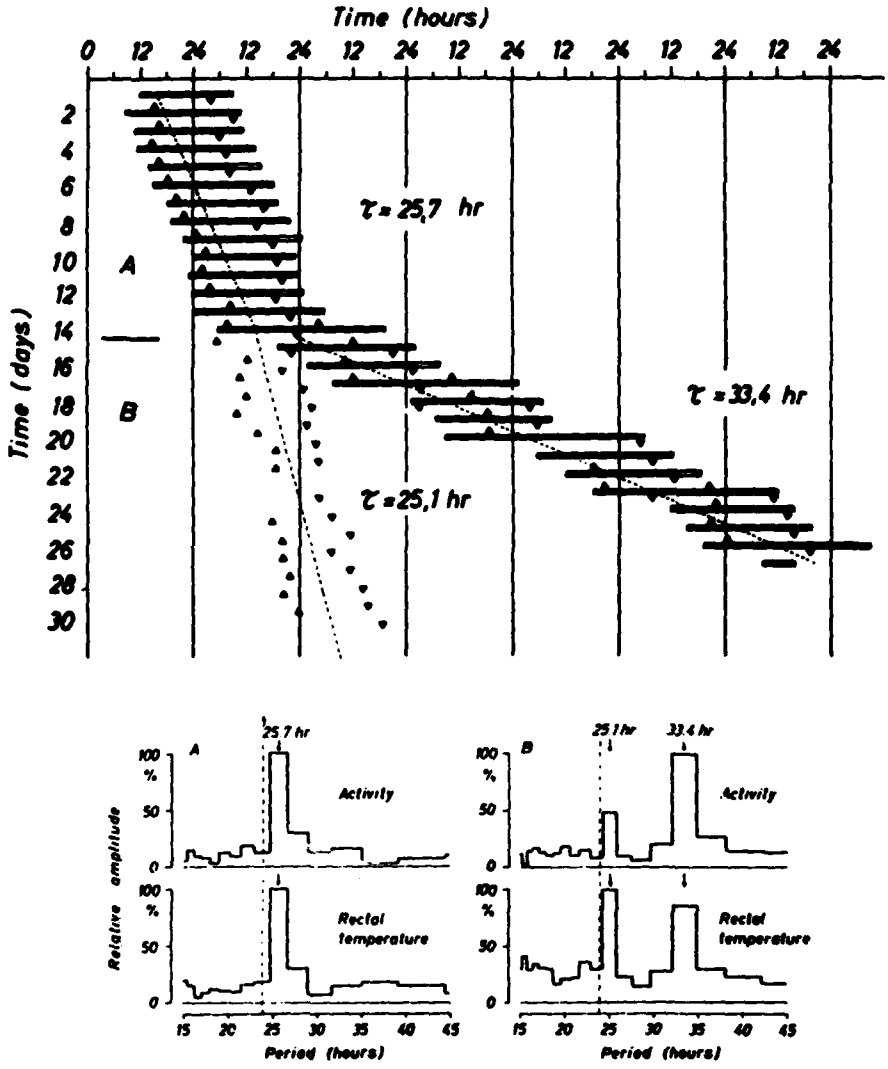


FIG. 1. Top: Activity rhythm (*solid bars*, activity; *open bars*, rest) and temperature rhythm (triangles show daily maxima and minima) of a human isolated from all time cues in an underground bunker. **A:** Internally synchronized free-running rhythms. **B:** Spontaneous internal desynchronization. **Bottom:** Periodograms from section A and section B. (From Wever, ref. 48, with permission.)

studies, desynchronization occurred between the renal rhythms, on the one hand, and the temperature and feeding rhythms, on the other. Although the activity rhythm was not measured, it would very likely follow the temperature and feeding rhythms. Thus, there was no evidence for a desynchronization

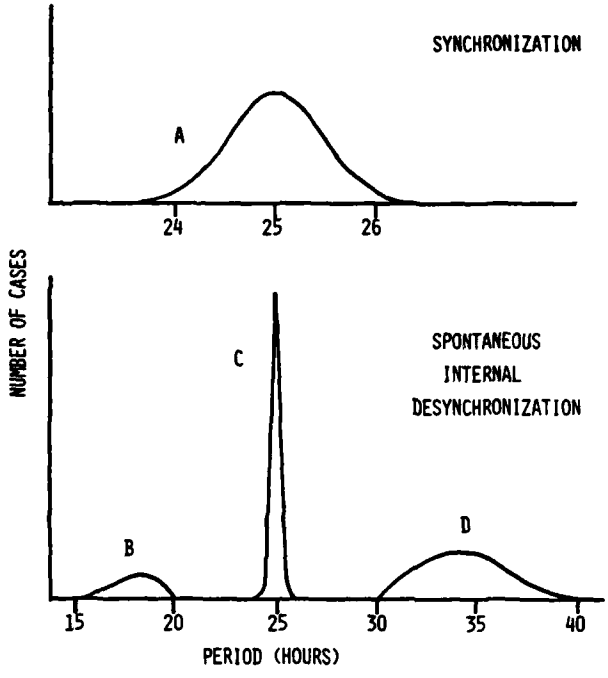


FIG. 2. Diagrammatic representation of all the major periodogram peaks produced by the human subjects who showed internally synchronized free-running rhythms and true spontaneous internal desynchronization when isolated in the underground bunkers by Aschoff's group. Apparent desynchronization is not included. Although these distributions resemble the histograms of Wever (ref. 49, Fig. 37; ref. 48, Figs. 4, 5), they are not the same. Wever separates the periods according to whether they are dominant in temperature or activity, i.e., on the basis of the tallest periodogram peak for each rhythm. In this figure the periods are separated according to the type of pattern, synchronization or desynchronization.

between the temperature and activity rhythms in the monkey. Furthermore, none of the rhythms showed the extremely long or short periods, of more than 30 or less than 20 hr, seen in the human (34,40,41).

THE TRADITIONAL MODEL OF SPONTANEOUS INTERNAL DESYNCHRONIZATION

Wever (48,49) has developed a comprehensive multiple-oscillator theory of human circadian rhythms based on the phenomenon of spontaneous internal desynchronization. Recently, a similar model has been elaborated by Kronauer (*this volume*, and 30). In these models, which will be referred to as "traditional," temperature and activity are controlled primarily by separate oscillators, although each oscillator exerts some control over all the rhythms. In

Wever's model these oscillators are termed group I and group II; in Kronauer's model they are termed *x* and *y*; for simplicity, we shall refer to them as the temperature and activity oscillators. According to the traditional models, the temperature and activity oscillators are coupled and free-run together during the synchronized free runs. Eventually these oscillators uncouple and free-run more or less independently, revealing their own natural frequencies. This uncoupling is the cause of internal desynchronization. When the two oscillators are mutually synchronized, they assume a compromise period between the periods they would assume when free-running independently. This compromise period is closer to the period of the temperature oscillator than to the period of the activity oscillator, because the temperature oscillator is stronger. (For example, in Fig. 1, the compromise period of 25.7 hr is closer to 25.1 than to 33.4 hr.)

In Wever's theory, there is a broad, normal distribution of activity oscillators ranging from less than 15 hr to more than 40 hr, depending on the individual subject, whereas the temperature oscillators of all the subjects have a period close to 25 hr, ranging from about 24 to 26 hr. Those subjects who have activity oscillators in the circadian range, between about 20 and 30 hr, never desynchronize. Their activity oscillators always remain coupled to their temperature oscillators. Only those subjects who have activity oscillators on the fringes of the distribution, less than about 20 hr and greater than about 30 hr, show internal desynchronization. For these subjects the disparity between the period of their activity oscillator and the period of their temperature oscillator is so great that the oscillators eventually uncouple during constant conditions. This theory explains the multimodal distribution of periods seen during spontaneous internal desynchronization (Fig. 2) and accounts for the finding that the majority of subjects never desynchronize. In Kronauer's theory, the activity oscillators start out with a period in the circadian range, but this gradually lengthens over days in isolation, reaching values of 30 hr or more. All subjects will eventually desynchronize when the disparity between the periods of the activity and temperature oscillators becomes too great.

Wever's theory has served over the years to help organize and analyze the complex data on human circadian rhythms, and Kronauer's model follows in this tradition. A major drawback of these models is that it is necessary to postulate that humans have activity oscillators with unusually long or short periods outside of the usual circadian range. These extreme periods are not seen in other animals, not even in the squirrel monkeys who are described as internally desynchronized (41). Even in "splitting," where two distinct periods temporarily appear in the activity rhythm, both periods are in the circadian range (37). Some authors have been skeptical that human circadian rhythms can be controlled by oscillators with such periods, and they have suggested that these periods might represent some type of artifact (2,7,31,32,40). Nevertheless, most of us have accepted the idea that in this respect the circadian system of humans differs from that of the other animals.

THE PHASE-SHIFT MODEL OF SPONTANEOUS INTERNAL DESYNCHRONIZATION

The phase-shift model is a single-oscillator model that is an alternative to the traditional model of spontaneous internal desynchronization. It represents the convergence of various observations about human circadian rhythms and various computer models of circadian oscillators. Some of these observations will be summarized briefly before the phase-shift model is described.

In most cases of spontaneous internal desynchronization the activity rhythm shows a scalloped pattern (Fig. 1, section B) that has been described as consisting of "phase jumps" (49) or "clusters" (15). Czeisler pointed out that the free-running period of each cluster is the same as the period displayed during the synchronized free run preceding the desynchronization (15). This raises the possibility that the activity rhythm continues to free-run during internal desynchronization with the same period as during synchronization, but with the addition of occasional phase shifts. Figure 3 shows a computer model used to help interpret the data on rats exposed to non-24-hr light–dark (LD) cycles (20–22). This model shows the pattern produced by a sine wave (representing a circadian oscillator) that free-ran until it reached a certain phase relationship to the LD cycle, at which point it was abruptly shifted; then it free-ran again until the same phase relationship was reached. Then it was shifted again, etc. For this particular model, the free-running period of the oscillator was 24 hr, but this period did not appear in the periodogram (Fig. 3, bottom). Instead, there was a period at 22 hr, corresponding to the period of the LD cycle, and a period at 25.2 hr, which can be considered a mathematical artifact in the sense that it does not correspond to an oscillator in the model. This model shows that when time-series analyses are used on nonstationary data, the periods indicated do not necessarily reveal the periods of the underlying oscillators. If internal desynchronization consists of an alternation between free running (with a circadian period) and phase shifting, then the unusually long or short (i.e., noncircadian) periods in the periodograms will be considered mathematical artifacts. For example, the long period of 33.4 hr in Fig. 1 will be an artifact analogous to the period of 25.2 hr in Fig. 3.

During synchronized free runs, sleep begins near the minimum of the daily temperature cycle (47,49) (Fig. 1, section A, or Fig. 7, top). This is the phase at which subjective alertness and performance efficiency are the lowest, the phase at which sleep-deprived subjects feel the most sleepy (1,14,25,29). During internal desynchronization, sleep usually begins near the temperature minimum as well. However, there are also many sleep episodes that begin at unusual phases of the temperature cycle (15,16,52,53). Sleep does not always occur near the temperature minimum, even though subjective alertness has also reached its minimum (e.g., ref. 15, Fig. 62).

During internal desynchronization, the length of a sleep episode depends on its phase relative to the temperature rhythm. Most wake times occur on the

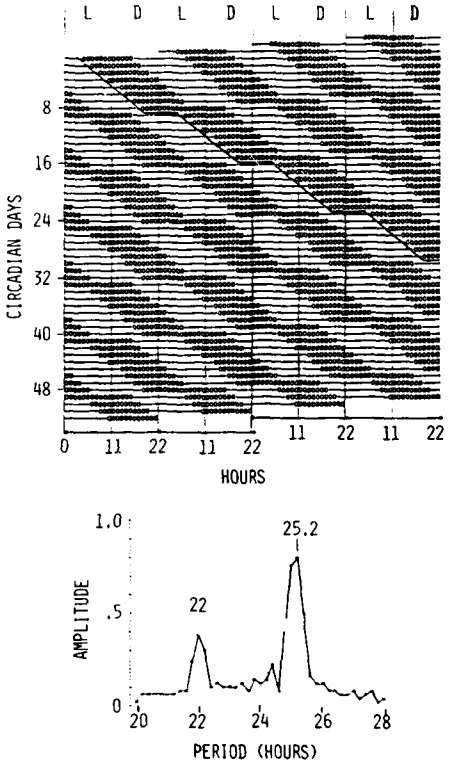


FIG. 3. Top: Computer model of a sine wave that free-ran until it reached a certain phase relationship to the 22-hr LD cycle. It was then abruptly shifted and then free-ran again until the same phase relationship was reached, etc. In the quadruple-plotted graph, each "O" represents an hour above the mean of each circadian day. **Bottom:** Periodogram (19) of the model data shown above. The dashed horizontal line represents the 95% confidence limit. (From Eastman, ref. 22, with permission.)

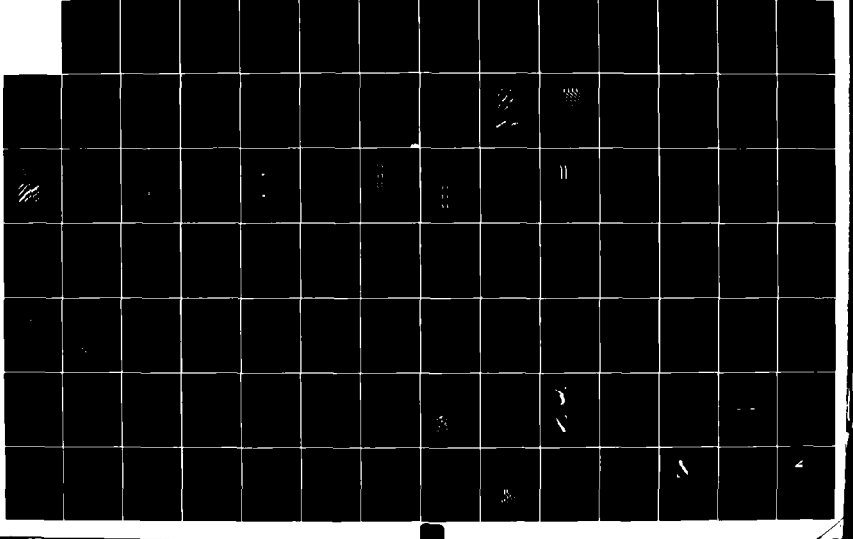
rising phase of the temperature cycle. Therefore, the longest sleep episodes are possible when sleep begins well before the rising phase. Accordingly, sleep episodes that begin near the temperature maximum are the longest (up to about 18 hr), sleep episodes that begin on the rising phase of the temperature cycle are the shortest, and sleep episodes that begin near the temperature minimum are of normal length (15,16,52,53). Czeisler's graphics (15) revealed that internal desynchronization consists of a repetitive pattern of long sleep episodes alternating with clusters of normal-length episodes.

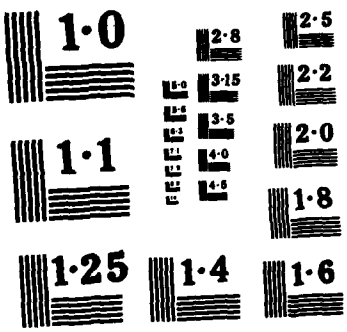
In addition to these recent findings that show how sleep and waking depend on the temperature rhythm, it is well known that sleep and waking affect body temperature through "masking" effects (4,10). Sleep lowers body temperature, and being awake and active raises body temperature. The magnitude of these masking effects depends on the circadian phase (33). Aschoff (6) emphasized this bidirectional mode of interaction between the sleep–wake rhythm and the body-temperature rhythm.

These and other observations led to the phase-shift model of internal desynchronization, as illustrated in Fig. 4. A single circadian oscillator controls both the temperature rhythm and the activity (sleep–wake) rhythm. For

AD A145 712 MATHEMATICAL MODELS OF THE CIRCADIAN SLEEP-WAKE CYCLE (U) HARVARD MEDICAL SCHOOL BOSTON MA DEPT OF PHYSIOLOGY AND BIOPHYSICS M C MOORE-EDE 01 MAY 84 2/3

UNCLASSIFIED AFOSR-TR-84-0781 AFOSR-81-0133 F/G 6/16 NL




1·0
1·1
1·25
2·8
3·15
3·5
4·0
4·5
2·5
2·2
2·0
1·8
1·4
1·6

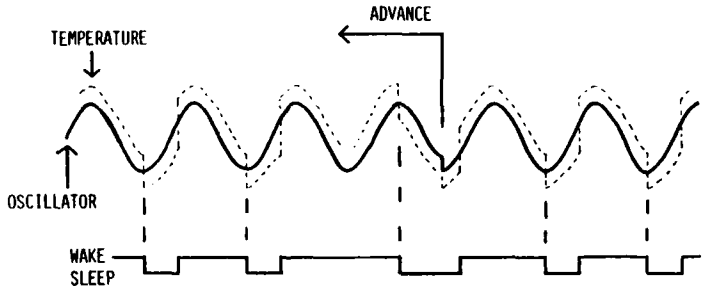


FIG. 4. Method for producing computer phase-shift models of spontaneous internal desynchronization. A single oscillator (*thick sine wave*) controls the temperature rhythm (*dotted line*) as well as the activity or sleep-wake rhythm (*square wave*). Masking is added to the temperature rhythm; temperature is raised during waking and lowered during sleep. Sleep usually begins on the minimum of the oscillation (in this case, sleep episodes 1, 2, 4, and 5), but sleep is occasionally misplaced (sleep episode 3). The misplaced sleep episode may have an unusual length, and it causes a small feedback shift (advance) in the circadian oscillator. (From Eastman, ref. 21, with permission.)

simplicity, the circadian oscillator is represented by a sine wave. Temperature follows the oscillator, but in addition masking is simulated; temperature is raised by a constant amount during waking and lowered by a constant amount during sleep. Sleep usually begins on the minimum of the temperature cycle. The signal to go to sleep, or the desire to go to sleep, might be a direct result of the falling body temperature. Alternatively, the signal could be transmitted from the circadian oscillator along other pathways, in which case the temperature cycle would merely serve as a convenient marker for the rhythm of sleepiness. Occasionally, sleep does not occur when temperature reaches its minimum. The subject may override the signal from the circadian oscillator for many reasons: he may want to finish an interesting book, or in his distorted sense of time it is not nighttime and he has been told to avoid naps, or he drank too much coffee, etc. The subject finally goes to sleep, perhaps because of the perception that enough time has passed and/or because of an accumulation of sleep "need." As a result, the onset of the subsequent sleep episode is misplaced, occurring later than the temperature minimum. This misplaced sleep episode will usually last until the next upward swing in body temperature and alertness (15,16,53) or, less frequently, may end early, perhaps because of hunger, some other discomfort, or sleep satiation.

In the phase-shift model, the misplaced sleep episode causes a small advance shift in the circadian oscillator (Fig. 4). In other words, the model proposes feedback from the activity (sleep–wake) rhythm to the circadian oscillator. Thus, there are actually two types of shifts in the phase-shift model, a shift in the activity rhythm because of the misplaced sleep episode and the subsequent shift

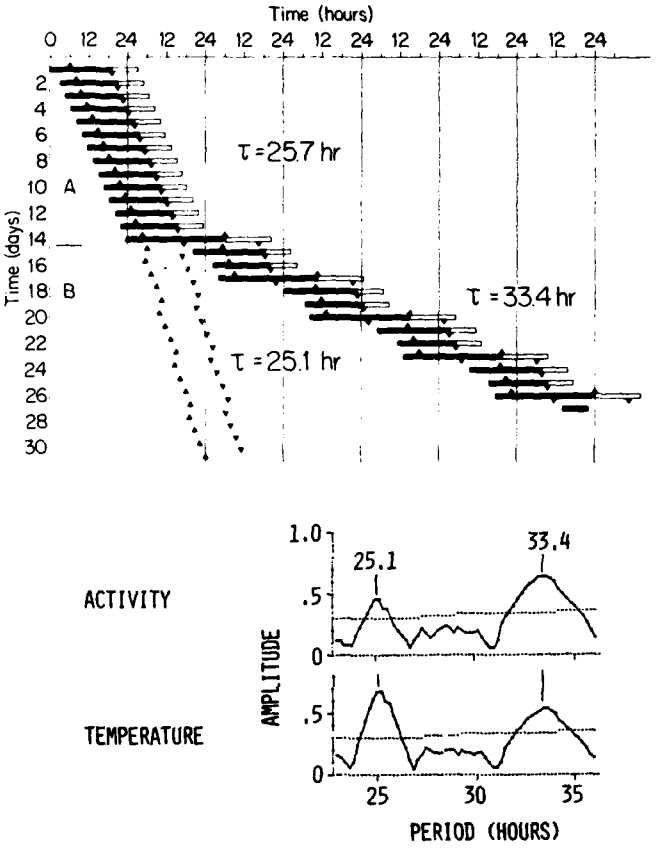


FIG. 5. Phase-shift model of the case of spontaneous internal desynchronization shown in Fig. 1 made by the method shown in Fig. 4. See text for parameters. **Top:** Symbols as in Fig. 1. **Bottom:** Periodograms (19) of the model data from section B.

in the circadian oscillator. The feedback from the activity rhythm to the circadian oscillator could be caused by sleep per se or, as suggested by A. Lewy (*personal communication*), could result from the light the subject is exposed to on awakening.

Variations of the simple method shown in Fig. 4 were used to mimic specific cases of spontaneous internal desynchronization. Figure 5 shows the phase-shift model designed to match the case of internal desynchronization shown in Fig. 1. For this model, the single circadian oscillator free-ran with a period of 25.7 hr throughout sections A and B. In section A, sleep episodes of 8 hr started on the minima of the oscillation. In section B, sleep onset skipped every fourth minimum, and then sleep began on the subsequent maximum. The mis-

placed sleep episodes were 14 hr long and advanced the circadian oscillator 2.34 hr. The sixth parameter specified the masking factor. Clearly, this model bears a close resemblance to the empirical data. The periodograms of the model are also a good match. The periodogram peaks reveal the same periods with similar relative heights, i.e., the 33.4-hr peak is taller for activity, and the 25.1-hr peak is taller for temperature.

Figure 6 shows the phase-shift model designed to match another case of spontaneous internal desynchronization. With this method of graphing, it is easy to see the pattern of long and normal-length sleep episodes during internal desynchronization that started at day 36. Eight long sleep episodes occurred out of phase with the clusters of normal-length sleep episodes. For this model, normal-length sleep episodes began on the minima of the circadian oscillation for about the first month. On day 36 the minimum was skipped; sleep finally began much later, and its duration was very long, causing the circadian oscillator to be shifted (advanced). Then skips and subsequent long sleeps recurred at more or less random intervals matched to the behavior of the subject. Notice how each cluster of normal-length sleep episodes is slightly advanced compared with the preceding cluster because of the advanced shifts of the circadian oscillator. The periodogram from the section of internal desynchronization (Fig. 6, bottom) has two main peaks. In the traditional model, the peak at 24.6 hr would be interpreted as the period of the temperature oscillator, and the peak at 29.2 hr would be interpreted as the period of the activity oscillator. In terms of the phase-shift model, all of the periodogram peaks are considered artifacts, because none corresponds to the period of the circadian oscillator, i.e., the period of the sine wave that was fed into the model, at 25.3 hr.

In conclusion, the phase-shift model demonstrates that spontaneous internal desynchronization can be produced while the circadian system is controlled by a single oscillator. It is not necessary to hypothesize separate temperature and activity oscillators. It is not necessary to hypothesize that activity is controlled by oscillators with periods outside of the usual circadian range. Finally, it is not necessary to hypothesize that spontaneous internal desynchronization is caused by an uncoupling or desynchronization between component oscillators.

THE PHASE-SHIFT MODEL: FURTHER DESCRIPTIONS AND SPECULATIONS

In the phase-shift model, spontaneous internal desynchronization is viewed as a sequence of normally placed and misplaced sleep episodes. Normally placed sleep episodes begin near the minima of the circadian temperature rhythm, and misplaced sleep episodes begin at other phases. The exact sequence of the normal and misplaced sleep episodes and the phase of the misplaced sleep episodes depend on the individual subject. The phase-shift

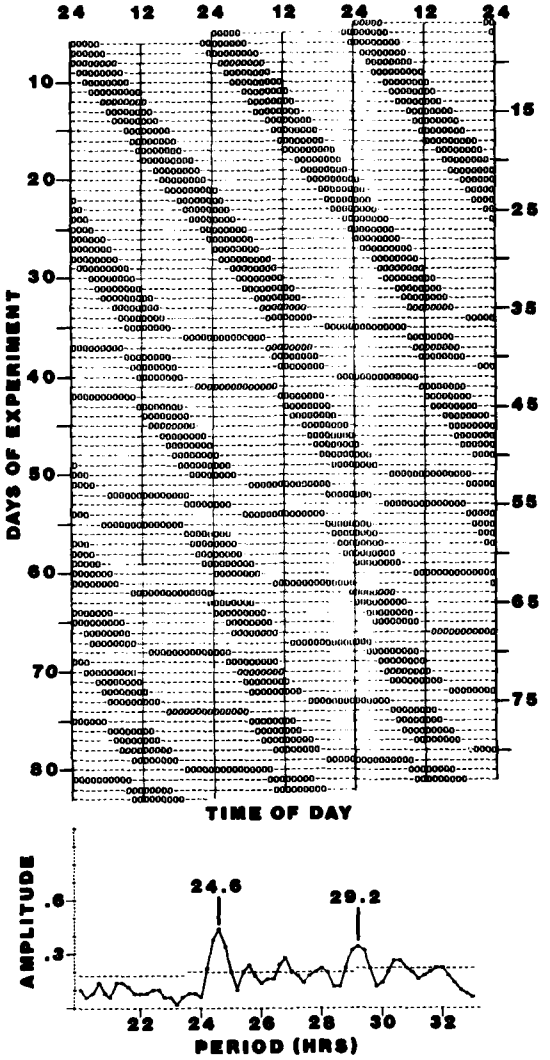


FIG. 6. Phase-shift model of a case of spontaneous internal desynchronization recorded by Czeisler et al. The original data can be found in the work of Kronauer (see Chapter 4), Czeisler et al. (16), Weitzman et al. (45), Kronauer et al. (30), or Czeisler (15). **Top:** Triple-plotted sleep chart; each O represents an hour of sleep. The circadian oscillator (*sine wave*) free-ran with a period of 25.3 hr. During the section of synchronized free-running (the first month), sleep episodes of 8 hr began on the minima of the oscillation. During the section of spontaneous internal desynchronization, various minima were skipped, so that sleep onset was delayed by 8.43 hr. The subsequent sleep episodes were long (14 hr) and advanced the circadian oscillator 4.5 hr. This model was made with all the parameters held constant except one, the days on which the skips occurred. **Bottom:** Periodogram (19) of the model data from days 35 to 74 of internal desynchronization. (From Eastman, ref. 22, with permission.)

models produced thus far are crude representations of a few individual patterns of spontaneous internal desynchronization. If this line of modeling were to be pursued, several refinements could be made. For example, the model might be improved by changing the constant masking factor of sleeping and waking on temperature to one that varies with the phase of the circadian cycle (33). For each specific model developed thus far, all the misplaced sleep episodes have started at exactly the same phase. More realistically, the phase of the misplaced sleep episodes for any individual subject probably varies somewhat. Phase-shift models can be made to mimic more complex patterns of internal desynchronization in which clear phase jumps or clusters are not seen. For example, the phase-shift model for one subject (not shown) consisted of two misplaced sleep episodes in a row, which appeared fairly regularly within a string of normally placed sleep episodes. Some subjects sleep, or at least take a nap, on almost every minimum, but in addition have some misplaced sleep episodes. Phase-shift models of these subjects would produce the extremely short periods between about 15 and 20 hr (Fig. 2).

The phase-shift model of internal desynchronization proposes a feedback shift from the activity rhythm to the circadian oscillator. This feedback shift was included not because of any direct evidence or preconceived notion that feedback should occur but merely because the shift was necessary to make the computer simulations match the empirical data. In the model, only the misplaced sleep episodes produce feedback, not the normally placed episodes that start near the minimum of the temperature cycle. If feedback from the activity rhythm to the circadian oscillator really exists, then it should occur in other situations, besides internal desynchronization, in which sleep does not occur near the minimum. Therefore, during entrainment to the 24-hr day, feedback should occur, because sleep begins about 90° before the temperature minimum (Fig. 7, middle) (47,49). It appears that the circadian oscillator is shifted during entrainment to the 24-hr day. There must be an advance shift of about 1 hr per day, because the average free-running period is about 25 hr. This advance could be due to feedback from the activity rhythm. Perhaps this placement of sleep several hours before the minimum helps advance the oscillator by the amount necessary for entrainment. In other words, one of the functions of sleep may be to help entrain the circadian oscillator. Alternatively, the perception of light on awakening may be the important factor in entrainment. During prolonged sleep deprivation there is a slight delay of circadian rhythms such as the temperature rhythm (5,6,49). These delays may result from the lack of normal sleep episodes, or normally timed awakenings, which could advance the circadian oscillator.

The phase-shift model for each subject was made by choosing several parameters based on a careful scrutiny of the subject's data and a trial-and-error process to better approximate the individual's pattern of desynchronization. Two of these parameters were the phase of the misplaced sleep episodes and the amount the circadian oscillator was shifted. A graph relating these two

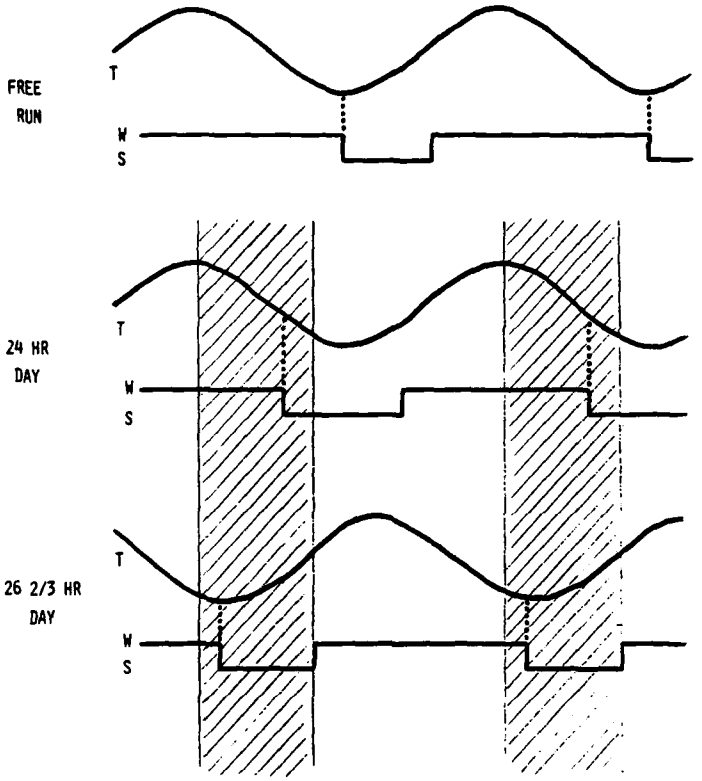


FIG. 7. Diagram of phase relationships during synchronized free runs (**top**), entrainment to a 24-hr day (**middle**), and *entrainment* to a 26⅔-hr day (**bottom**). The sine wave represents the temperature rhythm (T). The square wave represents the rhythm of sleeping (S) and waking (W). The shaded area represents the dark portion of the LD cycle. The dotted line emphasizes the phase at which the sleep episodes begin. The free-run diagram (**top**) is made from the average of many subjects, as shown by Wever (ref. 49, Fig. 17). The 24- and 26⅔-hr diagrams are redrawn from the data of a single subject isolated in an underground bunker. He was exposed to an overhead LD cycle with twilights and periodic gongs and could use reading lamps during the dark phase [subject 72a in Aschoff et al. (8); also known as subject S.R. in Wever (ref. 49, Fig. 69)]. This subject's data can also be seen in the work of Aschoff and Wever (ref. 10, Fig. 7B) and Kronauer (ref. 30, Fig. 10B). Other subjects show similar changes in internal phase relationships (8).

parameters from a few specific phase-shift models and the case of normal *entrainment* to the 24-hr day showed that the amount of shift appeared to be related to the phase of the misplaced sleep episode. In principle, it should be possible to generate a full phase response curve (PRC) using a large sample of phase-shift models of desynchronized subjects. In other words, the shifts of the

circadian oscillator might be shown to vary in magnitude and direction (advances or delays) depending on the phase of the sleep episodes, or of the awakening light.

The phase-shift model was the first to demonstrate that there are alternatives to the traditional model of spontaneous internal desynchronization. The strategy has been to show that the most typical, most publicized, individual cases of spontaneous internal desynchronization can be mimicked with a single-oscillator model. In the phase-shift model there are several determinants of the overt circadian rhythms, including the circadian oscillator, masking, "voluntary" behavior (behavior that cannot be predicted in these experiments), and a process of sleep need and recovery. However, the sleep-need-and-recovery factor was invoked in only the most rudimentary fashion. Daan and Beersma (*this volume*) have developed a more sophisticated single-oscillator model that includes a precise description of a sleep-need-and-recovery function based on Borbély's model (11). In contrast to the activity oscillator of the traditional model, this cycle of sleep need and recovery is a relaxation process. It is not controlled by a true self-sustained, fixed-period oscillator. In the more conventional style of modeling, a set of general solutions was produced by varying one parameter at a time. Their model can account fairly well for a wide range of human circadian phenomena. Feedback from the activity rhythm to the circadian oscillator was not included.

Earlier versions of the phase-shift model (20–22) proposed that a second shift-inducing oscillator, either internal or external to the subject, might be invoked to explain cases of internal desynchronization in which the misplaced sleep episodes occurred regularly (Fig. 1). There was no reason to postulate a second oscillator in cases with irregular patterns (Fig. 6). Daan and Beersma (*this volume*) have shown that regularly misplaced sleep episodes can arise from a single oscillator plus a cycle of sleep need and recovery. Their demonstration shows that there is little need for a shift-inducing oscillator.

In the phase-shift model, spontaneous internal desynchronization occurs when the subject ignores the signal from the circadian oscillator and does not go to sleep near the minimum. We might ask why some individuals become desynchronized while others do not, or, if all subjects eventually desynchronize, why some subjects become desynchronized before others. One possibility might be that some subjects become highly motivated to stay awake, for example, to continue work on some important project, and thus tend to ignore the internal signals to go to sleep. Another possibility is that some individuals have weaker circadian oscillators that produce oscillations of smaller amplitude, so that they tend to misjudge the right time for sleep. Daan and Beersma (*this volume*) include amplitude changes as an important part of their model. By gradually decreasing the amplitude of the circadian oscillation, they show how synchronized free running can abruptly change to internal desynchronization and that at the extreme a very small amplitude can produce a bicircadian pattern. This type of progressive change from synchronization to desynchronization to a bicir-

cadian pattern was first described by Czeisler (15). In the phase-shift model this change occurs when the minimum is skipped more and more frequently. In Kronauer's model (*this volume*; and 30) this change is accounted for by a gradual lengthening in the period of the activity oscillator. In Daan and Beersma's model, a gradual decrease in amplitude can account for the changing pattern of desynchronization. Daan and Beersma have also pointed out that the amplitude of circadian rhythms free-running in constant conditions is typically lower than during entrainment. This can explain why desynchronization occurs more frequently in constant conditions than in normal life.

The phase-shift model of spontaneous internal desynchronization does not refute the traditional model; it is merely offered as an alternative. Hopefully, some tests or experimental manipulations will be designed to distinguish between the two models. Meanwhile, there is one component of the phase-shift model that can be tested, the feedback shift from the activity rhythm to the circadian oscillator. Temporally isolated subjects showing synchronized free-running rhythms could be instructed to sleep at various phases of their temperature cycle, and the resulting course of the free run could be determined. In other words, misplaced sleep episodes could be experimentally produced to determine if they shift the circadian oscillator. This would be the equivalent of generating a PRC. The phase-shift model is a crude first step in the development of alternative models of internal desynchronization. Even if many of its propositions become obsolete, at least it has served to remind us that the traditional model is still open to question, and it has prompted the development of other single-oscillator models (S. Daan and D. Beersma, *this volume*).

CHANGES IN INTERNAL PHASE RELATIONSHIPS BETWEEN FREE RUNNING AND ENTRAINMENT

Although spontaneous internal desynchronization has been considered the best evidence for the theory of separate temperature and activity oscillators, other observations of human circadian rhythms have often been cited as further support for the traditional theory. These observations include the internal phase relationships between the temperature and activity rhythms—how these change between synchronized free runs, entrainment to the 24-hr day, and entrainment to Zeitgebers of other periods (8–10,47,49). Figure 7 schematically shows the phase relationships between the temperature and activity rhythms during three conditions. As discussed earlier, during synchronized free runs (top diagram), sleep begins around the minimum of the temperature cycle, whereas during entrainment to the 24-hr day (middle diagram), sleep begins about 90° before the minimum. The traditional model can easily account for this change in the internal phase relationships. According to the model, the temperature and activity oscillators are mutually synchronized during the free-running condition, but they assume different phase relationships to the Zeitgeber during entrainment because they have different characteristics (10,49).

The bottom diagram shows that both the temperature and activity rhythms show an advanced phase to the LD cycle when entrained to the 26⅔-hr day, as compared with entrainment to the 24-hr day. The maximum of the temperature rhythm advances from around "lights off" in the 24-hr day to many hours before "lights off" in the 26⅔-hr day. The activity rhythm also advances. This subject, who is a "night owl" in the 24-hr day, waking up around "noon," becomes an "early bird" in the 26⅔-hr day, waking at the crack of "dawn." These advances are predicted by oscillator theory, which states that as the period of the Zeitgeber increases, the phase of circadian rhythms advances (3,46). Notice, however, that the temperature rhythm advances more than the activity rhythm. This difference in the amount of advance results in different internal phase relationships between temperature and activity in the 26⅔-hr day as compared with the 24-hr day. Sleep begins near the minimum of the temperature cycle in the 26⅔-hr day, whereas it begins about 90° before the minimum in the 24-hr day. Once again, this change in the internal phase relationships can be explained by the traditional model in which the temperature oscillator advances more than the activity oscillator as the Zeitgeber is lengthened, because the temperature oscillator is stronger (49).

Let us explore one of the ways in which these changes in internal phase relationships might be explained with a single circadian oscillator. The synchronized free-running condition (Fig. 7, top) is easily accounted for by a single oscillator that drives the temperature rhythm and triggers sleep near its minimum. Given this mechanism, how can we explain the fact that sleep begins about 90° before the temperature minimum during entrainment to the 24-hr day (Fig. 7, middle)? We can assume that the phase of the circadian oscillator and the rhythms it drives, such as the temperature rhythm, is set by virtue of its entrainment to the 24-hr Zeitgeber. If sleep were to begin at the temperature minimum, when the subject feels the most sleepy, then sleep would begin too late; it would begin during the light phase and would occur entirely within the light phase. This subject is a "night owl"; the phase of his temperature rhythm is delayed. For other subjects, the temperature minimum occurs somewhat earlier, near the end of the dark period, but sleep triggered at the minimum will still occur too late. These late sleep times will not be adaptive ecologically or socially, in most circumstances.

We can account for the earlier onset of sleep in the 24-hr day by considering other factors. The subject may go to bed before the minimum because he knows it has been dark for a long while and believes it is the proper time for sleep. This could be called masking by the LD cycle, or behavior prompted by the knowledge of time. In addition, he has been awake all day, has built up some sleep need, and may feel tired. Finally, his temperature has started to drop; so he is beginning to feel sleepy. All these factors, circadian rhythm, masking, and sleep need, may contribute to his desire and ability to go to sleep before the minimum. In the 26⅔-hr day, the circadian oscillator and therefore the temperature rhythm assume a new entrained phase relationship to the

Zeitgeber. The temperature minimum occurs shortly after "lights out." At this point the subject is the most sleepy, and he knows it is nighttime. All factors converge to make it a perfect time for bed. In conclusion, only one oscillator is necessary to explain these various phase relationships if we assume that other factors in addition to the circadian oscillator can influence the scheduling of sleep and wake in the human.

FORCED INTERNAL DESYNCHRONIZATION

Although internal desynchronization may never occur spontaneously in some subjects, at least during the time in isolation, it can be reliably forced by Zeitgebers with periods near the limits of entrainment. Figure 8 shows a subject who was entrained to the 24-hr Zeitgeber, but became desynchronized (as defined by the presence of more than one period in the activity periodogram) when a 28-hr Zeitgeber was applied. According to the traditional theory, 28 hr was outside the range of entrainment for the temperature oscillator, so that it free-ran with a period of 24.8 hr. On the other hand, the activity oscillator had a larger range of entrainment, so that it remained entrained, or relatively entrained, to the 28-hr Zeitgeber (9,10,48,49).

Once again, let us explore one of the ways these results can be explained with a single circadian oscillator. This oscillator could drive the temperature rhythm and free-run in the 28-hr day because 28 hr is beyond the limit of entrainment. The activity rhythm, on the other hand, could have been partially controlled by the circadian oscillator and partially influenced directly by the LD cycle. The daily chart (Fig. 8) shows that the onsets of activity followed the free-running temperature rhythm. On each day, activity started a little before the temperature maximum. On the other hand, the onsets of rest did not follow the free-running rhythm, but tended to occur near the beginning of the dark phase. This is perfectly understandable, because this subject did not have access to reading lamps and had little choice but to go to bed when the overhead lights were off. In this view, then, the activity rhythm consists of two components (both of which appear in the activity periodogram), a free-running component produced by the circadian oscillator and a masking component due to the LD cycle. According to this single-oscillator theory, the range of entrainment of the activity rhythm was not larger than the range of entrainment of the temperature rhythm; it only appeared to be larger because of masking by the 28-hr LD cycle. In conclusion, only one oscillator is necessary to explain forced internal desynchronization if we assume that other factors in addition to the circadian oscillator can influence the activity rhythm of humans.

CONCLUSIONS

The traditional model of human circadian rhythms, developed over the years primarily by Wever (48,49), can explain many phenomena, including spon-

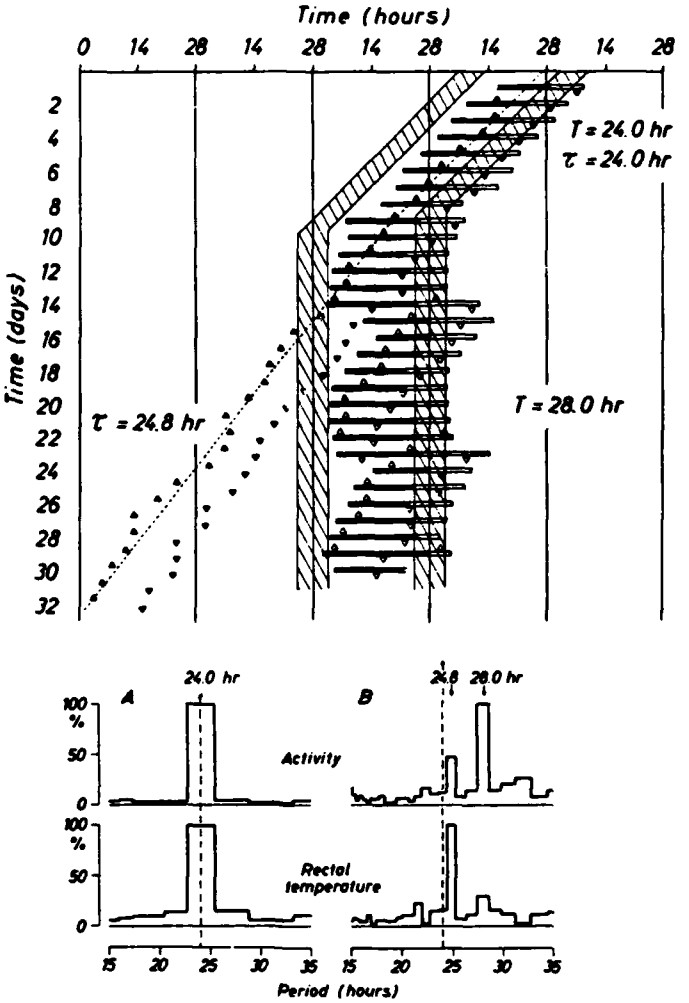


FIG. 8. Top: Daily chart of a subject isolated in an underground bunker and exposed to Zeitgebers consisting of an overhead LD cycle with twilights and periodic gongs. *Solid bars*, activity. *Open bars*, rest. *Triangles*, maxima and minima of the daily temperature cycle. *Shaded areas*, dark phase. For the first 8 days the subject was entrained to a 24-hr Zeitgeber. Then internal desynchronization was forced by a 28-hr Zeitgeber. Notice that this graph is plotted on a 28-hr time axis. **Bottom:** Periodograms during entrainment to the 24-hr day (*A*) and during forced internal desynchronization (*B*). (From Wever, ref. 49, with permission.)

taneous internal desynchronization, changes in internal phase relationships, and forced internal desynchronization. This theory relies on the concept of an activity oscillator that can have an extremely long or short free-running period and is weaker (and therefore has a larger range of entrainment) than the temperature oscillator. However, we have also seen that the same human circadian phenomena can also be explained by a single circadian oscillator. Rather than invoking a separate activity oscillator to explain the divergent pattern of activity, the single-oscillator model employs "voluntary" behavior, masking, and sleep need, in addition to the basic circadian oscillation.

In the excitement of unraveling the components of multiple-oscillator systems in recent years, we may have overestimated the need for multiple oscillators to explain much of the data. Any apparent dissociations or desynchronizations between rhythms are often ascribed to the uncoupling of component oscillators. We might learn more about circadian rhythms by considering how these patterns could be produced while the oscillators of the circadian system function synchronously as a single unit.

ACKNOWLEDGMENTS

This work was supported by N.I.M.H. Grant MH-4151 from the National Institutes of Health to Dr. Allan Rechtschaffen and by N.I.M.H. training Grant MH-14651 to the University of Chicago, Department of Psychiatry.

REFERENCES

1. Åkerstedt, T., Froberg, J. E., Friberg, Y., and Wetterberg, L. (1979): Melatonin excretion, body temperature and subjective arousal during 64 hours of sleep deprivation. *Psychoneuroendocrinology*, 4:219–225.
2. Aschoff, J. (1965): Circadian rhythms in man. *Science*, 148:1427–1432.
3. Aschoff, J. (1965): The phase-angle difference in circadian periodicity. In: *Circadian Clocks*, edited by J. Aschoff, pp. 262–276. Elsevier/North Holland, Amsterdam.
4. Aschoff, J. (1970): Circadian rhythm of activity and of body temperature. In: *Physiological and Behavioral Temperature Regulation*, edited by J. D. Hardy, A. P. Gagge, and J. A. J. Stolwijk, pp. 905–919. Charles C Thomas, Springfield, Ill.
5. Aschoff, J. (1978): Features of circadian rhythms relevant for the design of shift schedules. *Ergonomics*, 21:739–754.
6. Aschoff, J. (1981): Circadian rhythms interference with and dependence on work-rest schedules. In: *The 24 Hour Workday. A Symposium on Variations in Work-Sleep Schedules*, edited by L. C. Johnson, D. I. Tepas, W. P. Colquhoun, and M. J. Colligan, pp. 13–50. National Institute for Occupational Safety and Health, Washington, D.C.
7. Aschoff, J., Gerecke, U., and Wever, R. (1967): Desynchronization of human circadian rhythms. *Jpn. J. Physiol.*, 17:450–457.
8. Aschoff, J., Poppel, E., and Wever, R. (1969): Circadian rhythms in men as influenced by artificial light-dark cycles of various periods. *Pfluegers Arch.*, 306:58–70.
9. Aschoff, J., and Wever, R. (1976): Human circadian rhythms: A multioscillatory system. *Fed. Proc.*, 35:2326–2332.
10. Aschoff, J., and Wever, R. (1981): The circadian system of man. In: *Handbook of Behavioral Neurobiology, Vol. 4: Biological Rhythms*, edited by J. Aschoff, pp. 311–331. Plenum Press, New York.

11. Borbély, A. A. (1980): Sleep: Circadian rhythm versus recovery process. In: *Functional States of the Brain: Their Determinants*, edited by M. Koukkou, D. Lehmann, and J. Angst, pp. 151–161. Elsevier/North Holland, Amsterdam.
12. Boulos, Z., and Terman, M. (1980): Food availability and daily biological rhythms. *Neurosci. Biobehav. Rev.*, 4:119–131.
13. Bunning, E. (1973): *The Physiological Clock*. English Universities Press, London.
14. Colquhoun, W. P. (1971): Circadian variations in mental efficiency. In: *Biological Rhythms and Human Performance*, edited by W. P. Colquhoun, pp. 39–108. Academic Press, New York.
15. Czeisler, C. A. (1978): Human circadian physiology: Internal organization of temperature, sleep-wake and neuroendocrine rhythms monitored in an environment free of time cues. Ph.D. dissertation, Stanford University.
16. Czeisler, C. A., Weitzman, E. D., Moore-Ede, M. C., Zimmerman, J. C., and Knauer, R. S. (1980): Human sleep: Its duration and organization depend on its circadian phase. *Science*, 210:1264–1267.
17. Daan, S., and Berde, C. (1978): Two coupled oscillators: Simulations of the circadian pacemaker in mammalian activity rhythms. *J. Theoret. Biol.*, 70:297–313.
18. Davis, F. C., and Menaker, M. (1980): Hamsters through time's window: Temporal structure of hamster locomotor rhythmicity. *Am. J. Physiol.*, 239:R149–R155.
19. Dorrscheidt, G. J., and Beck, L. (1975): Advanced methods for evaluating characteristic parameters (τ,α,ρ) of circadian rhythms. *J. Math. Biol.*, 2:107–121.
20. Eastman, C. I. (1980): Circadian rhythms of temperature, waking, and activity in the rat: Dissociations, desynchronizations, and disintegrations. Ph.D. dissertation, University of Chicago.
21. Eastman, C. I. (1981): Circadian rhythms of rats at the limits of entrainment and a phase-shift model of "spontaneous internal desynchronization" in humans. In: *Sleep 1980, 5th European Congress on Sleep Research*, edited by W. P. Koella, pp. 59–64. S. Karger, Basel.
22. Eastman, C. I. (1982): The phase-shift model of spontaneous internal desynchronization in humans. In: *Vertebrate Circadian Systems: Structure and Physiology*, edited by J. Aschoff, S. Daan, and G. Groos, pp. 262–267. Springer-Verlag, New York.
23. Enright, J. T. (1980): Temporal precision in circadian systems: A reliable neuronal clock from unreliable components? *Science*, 209:1542–1544.
24. Enright, J. T. (1980): *The Timing of Sleep and Wakefulness*. Springer-Verlag, New York.
25. Gillberg, M., and Akerstedt, T. (1981): Possible measures of "sleepiness" for the evaluation of disturbed and displaced sleep. In: *Night and Shift Work: Biological and Social Aspects*, edited by A. Reinberg, pp. 155–160. Pergamon Press, London.
26. Hardeland, R., Hohmann, D., and Rensing, L. (1973): The rhythmic organization of rodent liver: A review. *J. Interdiscip. Cycle Res.*, 4:89–118.
27. Jouvet, M., Mouret, J., Chouvet, G., and Siffre, M. (1974): Toward a 48 hour day: Experimental bicircadian rhythm in man. In: *The Neuroscience Third Study Program*, edited by F. O. Schmitt and F. G. Worden, pp. 491–497. M.I.T. Press, Cambridge.
28. Kawato, M., and Suziki, R. (1980): Two coupled neural oscillators as a model of the circadian pacemaker. *J. Theor. Biol.*, 86:547–575.
29. Kleitman, N. (1963): *Sleep and Wakefulness*. University of Chicago Press, Chicago.
30. Kronauer, R. E., Czeisler, C. A., Pilato, S. F., Moore-Ede, M. C., and Weitzman, E. D. (1982): Mathematical model of the human circadian system with two interacting oscillators. *Am. J. Physiol.*, 242:R3–R17.
31. Mills, J. N. (1973): Transmission processes between clock and manifestations. In: *Biological Aspects of Circadian Rhythms*, edited by J. N. Mills, pp. 27–83. Plenum Press, New York.
32. Mills, J. N., Minors, D. S., and Waterhouse, J. M. (1974): The circadian rhythms of human subjects without timepieces or indication of the alteration of day and night. *J. Physiol.*, 240:567–594.
33. Mills, J. N., Minors, D. S., and Waterhouse, J. M. (1978): The effect of sleep upon human circadian rhythms. *Chronobiologia*, 5:14–27.
34. Moore-Ede, M. C., and Sulzman, F. M. (1981): Internal temporal order. In: *Handbook of Behavioral Neurobiology, Vol. 4, Biological Rhythms*, edited by J. Aschoff, pp. 215–241. Plenum Press, New York.

35. Pavlidis, T. (1969): Populations of interacting oscillations and circadian rhythms. *J. Theor. Biol.*, 22:418–436.
36. Pavlidis, T. (1973): *Biological Oscillators: Their Mathematical Analysis*. Academic Press, New York.
37. Pittendrigh, C. S., and Daan, S. (1976): A functional analysis of circadian pacemakers in nocturnal rodents. V. Pacemaker structure: A clock for all seasons. *J. Comp. Physiol.*, 106:333–355.
38. Rusak, B. (1979): Neural mechanisms for entrainment and generation of mammalian circadian rhythms. *Fed. Proc.*, 38:2589–2595.
39. Shiotsuka, R., Jovonovich, J., and Jovonovich, J. A. (1974): In vitro data on drug sensitivity: Circadian and ultradian corticosterone rhythms in adrenal organ cultures. In: *Chronobiological Aspects of Endocrinology*. edited by J. Aschoff, F. Ceresa, and F. Halberg, pp. 255–267. Schaffauer-Verlag, New York.
40. Sulzman, F. M., Fuller, C. A., and Moore-Ede, M. C. (1977): Spontaneous internal desynchronization of circadian rhythms in the squirrel monkey. *Comp. Biochem. Physiol.*, 58A:63–67.
41. Sulzman, F. M., Fuller, C. A., and Moore-Ede, M. C. (1979): Tonic effects of light on the circadian system of the squirrel monkey. *J. Comp. Physiol.*, 129:43–50.
42. Takahashi, J. S., and Menaker, M. (1979): Physiology of avian circadian pacemakers. *Fed. Proc.*, 38:2583–2588.
43. Tharp, G. D., and Folk, G. E., Jr. (1964): Rhythmic changes in rate of the mammalian heart and heart cells during prolonged isolation. *Comp. Biochem. Physiol.*, 14:255–273.
44. Webb, W. B., and Agnew, H. W. (1974): Sleep and waking in a time free environment. *Aerospace Med.*, 45:617–622.
45. Weitzman, E. D., Czeisler, C. A., and Moore-Ede, M. C. (1979): Sleep-wake, neuroendocrine and body temperature circadian rhythms under entrained and non-entrained (free-running) conditions in man. In: *Biological Rhythms and Their Central Mechanism*, edited by M. Suda, O. Hayaishi, and H. Nakagawa, pp. 199–227. Elsevier/North Holland, Amsterdam.
46. Wever, R. (1965): A mathematical model for circadian rhythms. In: *Circadian Clocks*, edited by J. Aschoff, pp. 47–63. Elsevier/North Holland, Amsterdam.
47. Wever, R. (1972): Mutual relations between different physiological functions in circadian rhythms in man. *J. Interdiscip. Cycle Res.*, 3:253–265.
48. Wever, R. (1975): The circadian multi-oscillator system of man. *Int. J. Chronobiol.*, 3:19–55.
49. Wever, R. (1979): *The Circadian System of Man*. Springer-Verlag, New York.
50. Winfree, A. T. (1967): Biological rhythms and the behavior of populations of coupled oscillators. *J. Theor. Biol.*, 16:15–42.
51. Winfree, A. T. (1980): *The Geometry of Biological Time*. Springer-Verlag, New York.
52. Zulley, J., and Wever, R. A. (1982): Interaction between the sleep-wake cycle and the rhythms of rectal temperature. In: *Vertebrate Circadian Systems: Structure and Physiology*, edited by J. Aschoff, S. Daan, and G. Groos, pp. 253–261. Springer-Verlag, New York.
53. Zulley, J., Wever, R., and Aschoff, J. (1981): The dependence of onset and duration of sleep on the circadian rhythm of rectal temperature. *Pfluegers Arch.*, 391:314–318.

DISCUSSION

Dr. Czeisler: Most of your phase-shift data seem to be based on free-running patterns of human subjects in which there is a reversion after each phase jump to the period seen during the internally synchronized free run. Have you modeled subjects who consistently have a 30- or 40-hr free-running period?

Dr. Eastman: The more often the subject skips the minimum, the longer the periods become, and you can get all the way up to bicircadian days if you skip every other

minimum. There are two patterns of bicircadian days. In one, the subject sleeps on a minimum, skips a minimum, sleeps on a minimum, etc., and you can explain that easily with this model. Another pattern is when the subject skips a minimum and then goes to sleep a little past the time where the temperature starts to drop, has a very long sleep, and then skips the next temperature minimum since it has not been that long since he has been up.

Dr. Czeisler: When you model these shifts, do you have to take the original data, program each shift into a computer, and then say, "Now, there is another shift"? If so, it would then become a question of whether you have a model or whether you are just telling the computer to reproduce the actual data.

Dr. Eastman: Well, the model predicts that the subject will usually go to sleep on the minimum, and it will soon, I hope, predict how much feedback there is. Of course, if the subject skips the minimum because of some reason that only he knows (for example, if he is reading a book), we could never predict that. So, of course, I have to copy that from the real data.

Dr. Czeisler: Dr. Wever said in his talk that he does not like two oscillators because six parameters must be defined, but at least those are defined at the beginning, before the program is run.

Dr. Eastman: My parameters are all set for the whole model. There is only one parameter that changes, and that is the day on which there is a shift.

Dr. Czeisler: But that is the key thing, the thing that needs to be explained.

Dr. Eastman: What if that happens because of the subject's behavior? You cannot hope to predict that. That is the cognitive factor that people want to explain.

Mr. Pilato: To what extent is the pattern of skipping sleep periods systematic, and what kind of mechanism would you use to explain it? Does the subject have to count every 4 days? Is it a sleep-debt system, and is that an oscillator?

Dr. Eastman: Well, I think whatever you have, you have to introduce noise in terms of the subject's behavior. Serge Daan has what I consider to be a beautiful, sophisticated version of my model. [See S. Daan and D. Beersma, *this volume.*]

Dr. Weitzman: I think the point that bothers me most is the concept of an unknowable, not unknown, but ***unknowable***, factor. That is, from a scientific point of view, not acceptable. You rule out the search for causality by saying that the subject just decided to do some particular action and that is his "behavior."

Dr. Eastman: In my view you have got two things going. You have got an oscillator which induces a tendency to go to sleep at the minimum, but which the subject can ignore to a certain extent.

Dr. Czeisler: In our studies we are in frequent contact with the subject and have asked ourselves why they stay up late on some nights. The first few times subjects stayed up past their temperature troughs and remained awake for an extended length of time, we searched for things that might be keeping the subject awake. Was he trying to finish a crossword puzzle, or was he trying to read a book, as you suggest? The remarkable thing is that in the vast majority of cases they were not working on anything in particular. They were not in the middle of a project that was keeping them awake, as we might have thought. We have the subjects on a video monitor, and we can sit there quietly and watch them. They may prepare lunch or do something else. It is very hard to imagine while watching them that something is directly keeping them up.

Dr. Eastman: But is it not true that when they skip sleep they still feel tired when that minimum comes along?

Dr. Czeisler: Yes, according to their subjective assessments of alertness.

Dr. Borbély: Dr. Czeisler, do you discourage the subjects from napping?

Dr. Czeisler: Yes. As you know, the instructions to the subject were that they could go to sleep at any time that they desired, but we asked the subjects not to nap. When they chose to go to sleep, we ask that it be for the night.

Dr. Kronauer: I would like to make a mathematical comment. Dr. Eastman, first I would like to compliment you on the two-oscillator model that you have introduced using two periodicities. It is an ingenious way to take two periods that are both close to circadian and end up with one that is quite far removed. But because the one-oscillator model involves a skip every 3 or 4 days, you must recognize that what comes out of that at peak periods in ratios are the ratios of low integers, like 4 to 3. In other words, basically what you are generating is one oscillation with something on top of it which is a type of super subharmonic, and I think that the data simply do not support the fact that these desynchronized periods are the ratios of low integers.

Dr. Wever: In one experiment there was a subject in our bunker who was free-running during the first 2 weeks and then became synchronized to a 24-hr Zeitgeber. He was the first case we ever had of unintended social cues. The Zeitgeber was the female technician who was taking out the urine samples and who did not follow our instructions to go to the locked outer chamber once a day at very irregular intervals, if possible when the subject was sleeping. Later on, I found that the technician had fallen a little bit in love with the subject before the experiment. She left some letters in the locked chamber at a regular time. [*Laughter*] So the subject went into his outer chamber every 10 min until he found that his urines were removed, and there on some days was a letter. He was not aware, of course, that his rhythm changed. That was the first case we really had of social synchronization.

Mathematical Models of the Circadian Sleep-Wake Cycle, edited by M. C. Moore-Ede and C. A. Czeisler. Raven Press, New York © 1984.

Modeling Principles for Human Circadian Rhythms

Richard E. Kronauer

Division of Applied Sciences, Harvard University, Cambridge, Massachusetts 02138

The objective of this chapter is to present a review of features of human circadian rhythms that must affect any attempt to model the human circadian system. It is unfortunate but true that direct physiological knowledge of this system is so rudimentary as to afford almost no basis for model construction. This means that models are merely mathematical constructs that serve to organize our view of system performance. One model has an advantage over another only if it matches more data features or is more economical in form.

ESSENTIAL MODEL COMPLEXITY

Perhaps the most important single experimental observation is that the sleep–wake cycle and the body-temperature rhythm of the free-running human can exhibit two different circadian periods at the same time (1). Furthermore, these periods do not appear to be related by any ratio of low integers, so that one oscillation cannot be supposed to derive from the other by any relatively simple frequency multiplication/demultiplication scheme. Within the family of models described by ordinary differential equations it is possible to generate two incommensurate oscillating frequencies with a third-order system (2), but only when special care is exercised. The ordinary third-order system generates frequencies that are rationally related. The mathematical distinction between these two situations is that in the former the limit set is a toroidal surface, whereas in the latter the limit set is a line that spirals about the torus and ultimately closes in a finite number of cycles. If one wishes to avoid mathematical oddities and still retain the differential equation form, it makes sense to advance to the fourth-order system. This also permits a kind of parity between the two rhythms in that each can be viewed as originating principally in its own second-order system, while mathematically identifiable mechanisms generate interactions between these two subsystems. Thus, whereas a general fourth-order system is intrinsically more complex than a third-order system, the reduced fourth-order system represented by two coupled second-order systems can actually be the more economical form. Furthermore, the idea that the full system is a conjunction of two oscillatory subsystems is given physiological

support by studies in which lesions destroying the suprachiasmatic nuclei (SCN) in squirrel monkeys disrupt the rest–activity cycle but leave a persisting body-temperature rhythm (3). Other experiments in rats have shown that knife cuts in the hypothalamus that isolate the SCN neurally do not abolish circadian rhythmicity within the SCN (4). Thus, the SCN appear to act as a discrete pacemaker, but other oscillating centers also exist in the organism. When the sleep–wake cycle and body-temperature rhythm show different circadian periods, the two subsystems are understood to have lost internal synchrony.

FUNDAMENTAL VARIABLES AND THEIR PHYSIOLOGICAL REPRESENTATIONS

Following the nomenclature introduced earlier (5), we shall refer to x and y as the fundamental variables of the oscillatory subsystems. In so doing it is not necessary that the system be describable by differential equations, although some of the evidence to be cited later strongly suggests that such a description is realistic. When the x and y rhythms display different periods, it is possible to identify these periods in a variety of physiological variables in human subjects. It would be ideal if two readily measurable quantities could be found that would directly represent one or the other of the primary rhythmic variables x and y. Two quantities usually suggested for these roles are core body temperature, T_c, and sleep–wake itself (which we shall call the variable SW; SW = 1 corresponds to sleep, SW = 0 corresponds to wake). From the data it is evident that T_c and SW each contain a mixture of the two rhythms, but T_c is dominated by the period close to 24 hr (which we call x), whereas SW is dominated by the other period (typically 30 hr or more, which we call y).

At this point it is necessary to address a very important functional distinction: intrinsic summation versus output summation. It is the same distinction that is drawn between feedback and feedforward in control theory. Suppose, for example, that the x subsystem affects a particular physiological variable such as plasma cortisol. If cortisol then affects the y subsystem, a mechanism is established for x to exert a drive onto y, or to couple x and y. Either the cortisol or some other variables can similarly couple y into x. Given the numerous variables that display strong circadian variation, it would be most surprising if significant coupling of both kinds did not exist in the human circadian timing system. Consequently, even within the subsystems there would always be a mixture of the two rhythms. Neither x nor y will display a pure periodicity. The mere fact that a physiological variable exhibits both rhythms is no grounds for rejecting that variable as being a possible "perfect" representation of either x or y. On the other hand, a quite different situation can be envisaged. Suppose that the coupling (internal feedback) is weak, so that x and y are almost pure rhythmic variables (only slightly contaminated). It is possible that some variable other than x or y can receive strong drives from both subsystems and yet have little effect back on the subsystems. This third variable is in a

feedforward position with respect to x and y. It can also be said to combine or sum the outputs of the subsystems.

Using the foregoing definitions, when the system is internally desynchronized, the dominant period in T_c is that of x, whereas the dominant period in SW is that of y. The question remains how well T_c represents x and SW represents y. The previous discussion shows the question to be essentially unanswerable unless interventions can be made that either isolate the subsystems or drastically alter the coupling between them.

The only way this unanswerable question can be approached at present is through arguments of plausibility and consistency, based on additional data. Individual subjects are often observed to progress from internal synchrony to desynchrony during prolonged temporal isolation. During synchrony the composite rhythm has a period lying between the periods that T_c and SW show during desynchrony, and this period is heavily biased in favor of T_c (6). The implication is that the drive of x onto y is much stronger than the drive of y onto x [a ratio of about 4, according to a model of coupled Van der Pol oscillators (5)]. A measure of the overall strength of the interoscillator coupling is afforded by the differences of x and y periods when desynchrony first appears (typically an 18% to 30% difference).

Putting these observations together, we are led to expect that the intrinsic y variable will necessarily contain a significant component with the period of x, while x itself will remain relatively pure. It is therefore highly improbable that any observed rhythmic variable will have only the y periodicity (unless, by perverse chance, output summation puts just the correct amount of negative x into some variable so as to cancel the component of x already within y). Thus, SW becomes a plausible candidate for y itself (insofar as timing is concerned), and some simulations support this hypothesis (5). On the other hand, the fact that T_c contains a significant component with the y period indicates output summation and suggests that T_c is not a straightforward indicator of x. Another way in which this output summation is characterized is to say that there is a component of core temperature that is "evoked" by activity or suppressed by sleep (7). This concept is quite widely accepted, but it is important to note that hidden within it is the assumption that the drive from y onto the x subsystem is not important in this context.

THE CHARACTER OF SYNCHRONY LOSS

A very important feature of human free-run data is that synchrony and desynchrony are not the only states observed. A condition commonly exists in which the SW rhythm has, on average, the same period as the T_c rhythm, but SW has large excursions of phase with respect to T_c. Although some of these phase variations are undoubtedly random, there are also significant regular components that modulate the phase with periods in the range of 4 to 7 days. This phenomenon is called phase trapping (5). Three examples are shown in

Fig. 1. The middle of each sleep episode is marked, and a slightly smoothed line is drawn through or near these points to emphasize the fundamental regularities. Observe how in each case most of the variations of phase can be accommodated by this regular modulation. Note that the modulation periods are comparable to the beat periods between x and y that are observed early in desynchrony (see Figs 2 and 4A). Phase trapping is often encountered as an intermediate stage between synchrony and desynchrony; this suggests that phase trapping and desynchrony are related and that desynchrony may be the further development of a process begun with the initiation of phase trapping. Phase trapping actually represents the failure of the free-running system to enforce synchrony between x and y, and in this sense phase trapping bears a closer relation to desynchrony than to synchrony.

There is a simple explanation of phase trapping within the context of quasilinear oscillators that, although specialized, is worth summarizing here, because it may possibly have more universal implications. The explanation begins with the idea that there is a progressive reduction of the tendency of x and y to synchronize, either because of a reduction in coupling strength or because of an increase in the disparity of the oscillatory periods intrinsic to the x and y subsystems. (I shall offer evidence to suggest that it is more likely the increase in disparity of period that reduces the tendency to synchronize.) At some point, synchrony is lost, and the y subsystem begins to develop a small component at or near it intrinsic period, in addition to the large component already imposed on it by x. Further reduction of the tendency to synchronize causes progressive growth of the intrinsic component, with concomitant growth in the amplitude of phase-trapped modulations. All the while the average period is that of x, because the imposed x component is still the larger. The transition to desynchrony occurs when the intrinsic component becomes larger than the imposed component, whereupon the average period of y becomes that of the intrinsic component.

This explanation is extremely economical in terms of modeling. Furthermore, phase trapping appears to be very difficult to explain in any other way. In any event, phase trapping is a phenomenon that any credible model must be able to accommodate. A consequence of this interpretation of the relationship between phase trapping and desynchrony is that the transition from the former to the latter is the result of a minor change in the relative strengths of imposed and intrinsic components. The overt effect of the transition is a dramatic lengthening of the average period of SW (typically to 30 hr from a previous value of 25 hr), and usually with an exceptionally long subjective day–night cycle at the very beginning (often 40 hr). It is also possible to imagine that in some subjects the circadian system might remain poised at or very close to the transition threshold and under small extraneous influences might cross and recross the transition boundary. An example of this kind of system behavior is shown in Fig. 2. This subject lived in isolation from temporal cues for 170 days at the Laboratory of Human Chronophysiology, Montefiore Hospital, New York City. The subject

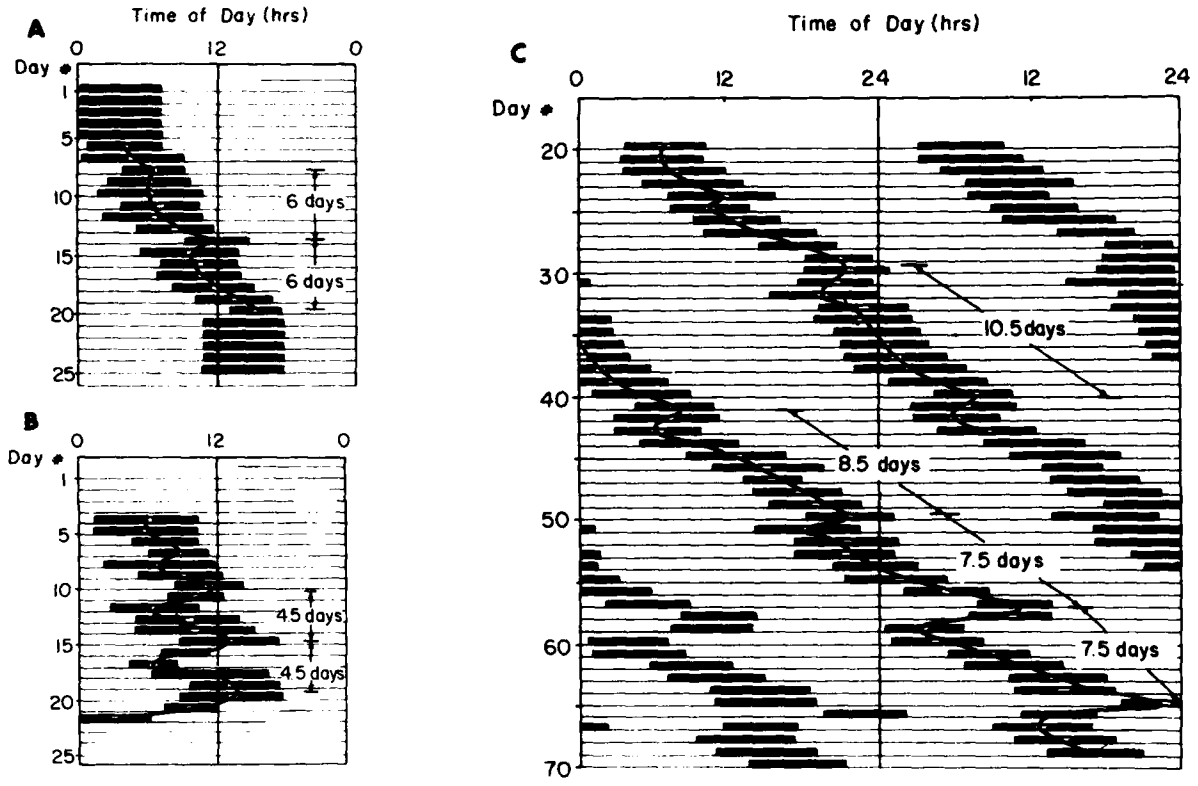


FIG. 1. A: Sleep episodes for free-running subjects showing phase trapping. *Solid bars*, polygraphically verified sleep (or bed rest, in the case of B). The center of each sleep episode is denoted by an open square, and a faired line has been drawn by eye to emphasize the underlying regular phase pattern. These data were reported in Fig. 14, subject FR07, of Czeisler (7). B: These data are adapted from Fig. 28 of Wever (6) and constitute the first 19 free-running bed-rest episodes. For the 7 days succeeding the data shown, the subject exhibited a complex desynchronization pattern, with rest episodes occurring, on average, once every 16 hr. C: These data correspond to days 35 through 70 of the full record reported in Fig. 2. As the full record shows, the subject first demonstrated internal desynchrony on day 82. See text for further discussion of the sleep pattern after day 82.

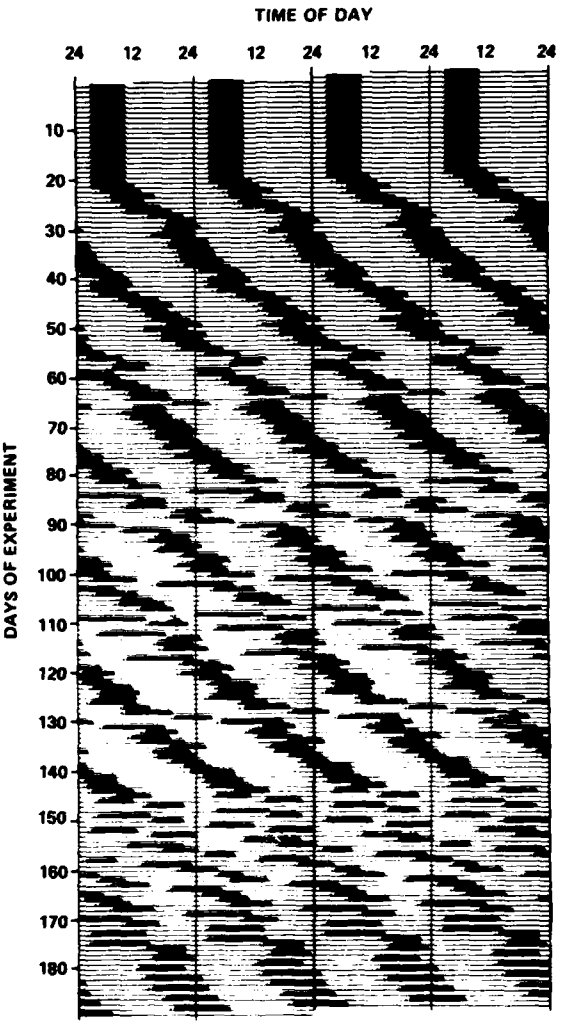


FIG. 2. Activity-rest cycle pattern plot: The sleep record of a subject allowed to run free for 170 days. These data were taken at the Laboratory of Human Chronophysiology, Montefiore Hospital, New York City, by C. A. Czeisler, J. M. Zimmerman, and E. D. Weitzman. Internal desynchronization occurs first at day 82, but the subject returns from time to time to the phase-trapped state. After day 145, the subject remains permanently in the state of internal desynchronization. The average length of the activity cycle increases progressively (but not monotonically) throughout the 6-month study. (*From Kronauer et al., ref. 10, with permission.*)

first became desynchronized on clock day 82 and remained so for 10 days. From day 92 to day 101 there were two phase-trapped modulation cycles, whereupon desynchrony resumed. Phase trapping reappeared between day 119 and day 146, with the exception of one cycle of desynchrony centered at day 130. After day 146 the subject remained desynchronized for the remainder of the experiment.

MODE OF ACTION OF A LIGHT/DARK ZEITGEBER

Further information about the circadian system has been provided by experiments in which external Zeitgebers have been manipulated. At the Montefiore Hospital facility, an imposed 24-hr light/dark cycle has been found to be an effective entrainer of SW in humans. In certain experiments of Aschoff and Wever (6), light/dark cycles alone were not effective in entraining SW at a variety of periods. The difference may well be due to the strictness of the light/dark cycle at Montefiore; no auxiliary light could be switched on by the subject during the dark (an available option in the relevant Aschoff and Wever cases). At other times, a stronger Zeitgeber protocol was also employed by Aschoff and Wever; this included the ringing of bells at intervals of one-eighth of the imposed period and a somewhat stricter observance of "dark." This stronger Zeitgeber entrained subject activity from periods as short as 18 hr to periods as long as 30 hr.

An important question is the mode of action of the light/dark stimulus (designated z) on the circadian system. In particular, does z act directly on one or both of the x and y subsystems? Because of the physiological evidence of direct input to the SCN from the retina via the retinohypothalamic tract in mammals (8), it is highly likely that z acts on the y subsystem. Because y acts on x, the pathway $z \rightarrow y \rightarrow x$ provides a two-stage connection from z to x. That is, the intrinsic coupling between y and x transmits the effect of z to x as well as to y. The same can be said of any possible direct action of z on the x subsystem; the action will be felt on y as well.

One of the amply documented (6,7) effects seen on release of a subject from a 24-hr entraining Zeitgeber is an adjustment of the relative phase of T_c and SW: T_c advances with respect to SW. The extent to which x advances with respect to y cannot be estimated with high accuracy because of the evoked effect of SW on T_c, but it is probably about 5 hr, typically. The substantial size of this shift (75° of phase) is very revealing. The phase relationship observed in free run represents the situation where x is effectively controlling y (because the drive of x onto y is so much stronger than that of y onto x). Under Zeitgeber action, if z were to enter the $x \rightleftarrows y$ feedback loop predominantly by drive onto x, the phase relation between x and y would have to be essentially the same as in free run. The large phase shift actually observed shows that z enters the loop predominantly by drive onto y (and perhaps exclusively so).

Substantiation of this view of the way in which z acts on the $x \rightleftarrows y$ system is afforded by phase-shift experiments. When subjects are transported across several time zones, the adjustment of the SW rhythm is effected rapidly (within a day or two). The adjustment of T_c proceeds much more slowly, seldom attaining a shift rate of 1 hr per clock day (9). If z were to act directly on x to a significant degree, the adjustment of T_c would occur at a rate comparable to SW. In fact, the slow adjustment of T_c reflects not only the absence of significant $z \rightarrow x$ drive but also the relative weakness of the $y \rightarrow x$ drive.

If one makes the conjecture that the "absence of significant $z \rightarrow x$ drive" is actually the total absence of drive from z to x (when z is a light/dark stimulus), then x is remarkably isolated from z. How then is one to explain the observation that, under appropriate conditions, such a Zeitgeber can entrain x without entraining y [as in Fig. 81 of Wever (6), reproduced here as Fig. 3]. This apparent paradox has a rational explanation and is also susceptible to modeling, as I shall show later. It is first necessary to point out that such a subject, in the absence of z, is in desynchrony and that the free-running periods of x and y are very disparate. Furthermore, the intrinsic period of x is within 1% or 2% of the period of the imposed z. Thus, in order for z to entrain x but not y, it is only necessary for the component that z induces in y to be insufficient to bridge the large disparity in z and y periods yet sufficient to bridge the tiny disparity in z and x periods (despite the large attenuation that component suffers through the

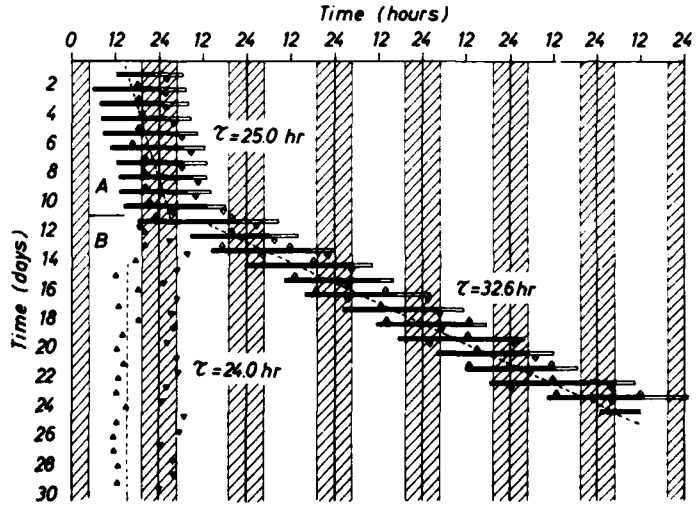


FIG. 3. A record of rest (*open bars*), activity (*filled bars*), body-temperature maxima (*upward-pointing triangles*), and temperature minima (*downward-pointing triangles*) for a subject under a light/dark Zeitgeber influence. (*Vertical hatched bars*, Zeitgeber dark.) (From Wever, ref. 6, with permission.)

weak $y \rightarrow x$ drive). In fact, there must always be a small band of disparities between the z and x periods for which such an entrainment will be found, regardless of whether or not y is entrained.

The converse situation to the one just described, where z entrains y but not x, hardly needs comment. All that is required is for the z and x periods to be sufficiently disparate that the combined effects of z and y (now both having the period of z) felt through the $y \rightarrow x$ coupling be insufficient to enforce synchrony of x. What is perhaps less obvious is that in either situation, where only one of the two subsystems is entrained to z, the entrainment of that subsystem is imperfect. Because the nonentrained subsystem possesses an average period different from that of z, and because the two subsystems are mutually coupled, it is inescapable that the entrained subsystem will contain a component having that different period. The component will be small, but nevertheless it will modulate the entrained variable so as to give the appearance of phase trapping with respect to z, rather than perfect synchrony with z.

Finally, two other situations can arise under the action of z. In one of these situations, x and y can proceed with the same average period, but it is a period that is different from that of z. (This is the case of a subject whose free-running state is either synchrony or phase trapping and for which the Zeitgeber either is very weak or has a large period disparity with the $x \rightleftarrows y$ composite.) In the other situation, x, y, and z all exhibit different periods. In either of these two situations there will be phase modulations of both x and y.

AN EXEMPLARY MODEL

My colleagues and I have presented elsewhere (5,10) a mathematical model that incorporates all of the performance features described. A brief summary will be given here, but for full details the earlier presentations should be consulted. Each of the x and y subsystems is represented by a Van der Pol oscillator, and the mutual coupling is represented by a single linear term in each oscillator. The Zeitgeber, z, is presumed to act on y alone. In nondimensional form, the model differential equations are

$$k^2\ddot{x} + k\mu_x(-1 + x^2)\dot{x} + \omega_x^2 x + F_{yx}k\dot{y} = 0$$

$$(k = \pi/12)$$

$$k^2\ddot{y} + k\mu_y(-1 + y^2)\dot{y} + \omega_y^2 y + F_{xy}k\dot{x} = F_{zy}z$$

As written, the intrinsic oscillator frequencies ω_x and ω_y assume values of unity if the corresponding periods are 24 hr. Periods longer than 24 hr are represented by ω_x or ω_y less than unity. The reasons for the particular form chosen for the mutual coupling will not be reviewed here, except to note that this form gives a compromise period for the two oscillators, when they synchronize, that lies between the intrinsic periods of the separate oscillators.

Parameter values that give good agreement with observed system performance are

$$F_{xy} = -0.16, \quad F_{yx} = -0.04, \quad \mu_x = 0.1, \quad \mu_y = 0.1,$$
$$\omega_x = 0.985, \quad 0.95 > \omega_y > 0.5$$

Because both μ_x and μ_y are much less than unity, the individual oscillators are of the quasilinear type (their limit cycles are almost circular, and their waveforms are almost pure sinusoids). For each, the transient adjustment in amplitude after an impulsive disturbance is relatively slow (approximately 10 radians or 2 cycles as the adjustment time constant). The ratio of the mutual coupling coefficients is 4, giving x a large dominance over y. Consequently, it is the absolute size of F_{xy} that establishes the range of frequency disparity, $\omega_x - \omega_y$, for which synchrony can be found. Approximately, when $|\omega_x - \omega_y| > |F_{xy}|$, phase trapping or desynchrony will be encountered. The intrinsic period assigned to the x oscillator is slightly over 24 hr (24,36); phase trapping is found when the period of the y oscillator (τ_y) exceeds 27.7 hr, and desynchrony occurs when τ_y is more than 29.3 hr.

As discussed elsewhere (5), studies of subjects in extended free run have shown that once desynchronization occurs, there is often a progressive lengthening of the period of SW. Figure 2 is a case in point: Between day 80 and day 140 the desynchronized sections show τ_y in the range 30 to 35 hr, whereas after day 140 the τ_y lie in the range 32 to 49 hr. The lengthening of τ_y is certainly not monotonic, but the overall trend is unmistakable. I postulate that a subject who displays internal synchrony when released into free run and then subsequently displays phase trapping and desynchrony does so because ω_y is progressively decreasing during these events. Other subjects who go directly into phase trapping or desynchrony on release presumably do so because ω_y is already sufficiently small when free run is begun. I have found that free-run records generally can be well simulated by our model while holding all parameters other than ω_y fixed throughout. I am consequently led to postulate that, for unexplained reasons, ω_y is a very labile parameter, whereas all the other system parameters are very stable over the time course of an experiment. Figure 4 shows a comparison, taken from an earlier study (5), between an extended free-run experiment and a model simulation in which ω_y was made to decline steadily while all other parameters remained fixed. There is reasonable qualitative agreement. In Fig. 4B (and in Figs. 7, 8, and 9), the sleep episodes (solid bars) correspond to the central two-thirds of the times when y is below zero, and the stippling corresponds to the times when x is below zero. Because of the evoked effects of SW on T_c (as yet inadequately understood), the episodes where T_c is below its mean should occupy some intermediate position within the stippled band and the sleep bars, presumably favoring the stippled portion.

Computer simulations reported earlier (5) show that there are relatively

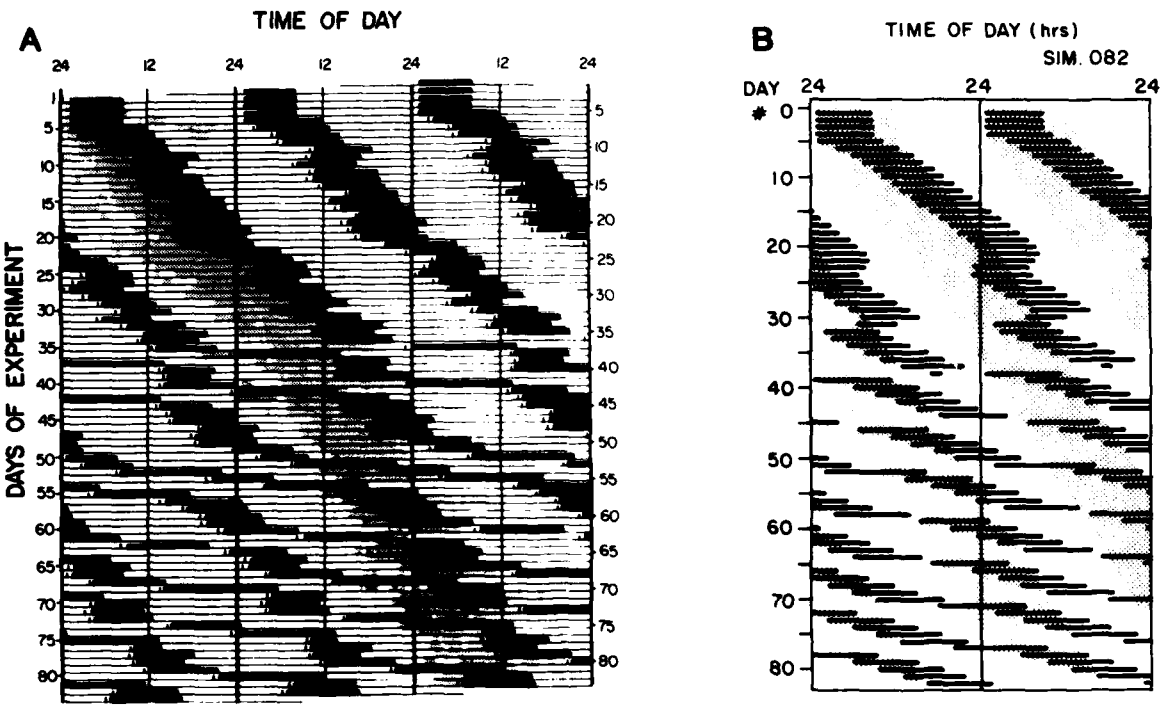


FIG. 4. A: Sleep episodes (*solid bars*) and body temperature below mean (*stippling*) for a subject released into free run at day 5. Prior to day 35 there is evidence of phase trapping, with modulation period in the range from 7 to 5 days. After day 35 the subject is internally desynchronized, with a phase-cluster modulation in the range from 5 to 4 days. Between days 43 and 52 it is possible that a phase-trapped cycle occurred. **B:** A computer simulation based on a model of coupled Van der Pol oscillators designed to mimic the data of **A**. Parameters are given in the text. *Solid bars* (sleep), central two-thirds of the times when y is below zero. *Stippling*, times when x is below zero. This is an approximation to T_c below its mean if sleep-evoked temperature effects are small. (From Kronauer et al., ref. 5, with permission.)

stringent limitations on model parameters for phase trapping to exist. In addition to the frequency disparity, $\omega_x - \omega_y$, having to be sufficiently large (but not too large, or desynchrony will occur), there are two other important conditions. First, μ_y must be quite small (less than 0.2 approximately), and, second, the drive from $x \rightarrow y$ must be considerably larger than the drive from $y \rightarrow x$. Both of these conditions obtain in the normal human circadian system, and it is especially interesting to note that the meeting of these conditions can be deduced from evidence that has nothing to do with the existence of phase trapping itself. The fact that μ_y is approximately 0.1 can be estimated from the time course of the internal phase adjustment that is seen on release from z. The much larger $x \rightarrow y$ drive than $y \rightarrow x$ drive is evidenced by the strong bias of the synchronized compromise period in favor of the intrinsic x period.

THE MODEL WITH PERIODIC ZEITGEBER EXCITATION

The inclusion of Zeitgeber drive presents the modeler with two fundamental questions: Is the waveform of z important? What amplitude of drive should be chosen to simulate the relatively weak or relatively strong light/dark stimuli described earlier? If z is periodic in the circadian range, and if the parameters of the model are in the normal range (close to those listed earlier), the waveform of z appears to be irrelevant for phenomena that involve two or more circadian cycles. Different waveforms are approximately equal in their effects, provided the fundamental Fourier component (i.e., the circadian component) is the same in the two waveforms. The reason is that the y subsystem is a quasilinear oscillator and acts like a resonant filter. Only the z component in that resonant range has any significant influence on y. Consequently, it is possible to use the simple functional form

$$z = F_{zy} \cos(k\omega_z t + \theta_z)$$

in entrainment studies. For acute phenomena, such as determining the length of rebound sleep after a specific sleep-deprivation episode, a more detailed description of z may be required.

How large the influence coefficient, F_{zy}, should be to represent weak or strong Zeitgebers can be estimated by choosing values that produce in the simulation entrainment limits that match observed limits. The strong Zeitgeber of Aschoff and Wever (6), for which entrainment of SW in the period range 20 to 30 hr has been demonstrated, requires F_{zy} to be approximately 2. At the other end of the range, a Zeitgeber that fails to entrain SW to an imposed 24-hr period when the intrinsic ω_y is 0.85 (a period of 28.5 hr) corresponds to F_{zy} of about 0.7. The entrainment of y also depends on the value assigned to ω_x, and the foregoing estimates are based on $\omega_x = 0.985$.

I have undertaken a study of entrainment limits for a fixed F_{zy} in which each of ω_x, ω_y, and ω_z take on a wide range of values relevant to experimental

conditions. All the other system parameters are held fixed at the values described earlier. Ostensibly, such a study involves three system parameters and is consequently very laborious. However, mathematical arguments (which will not be expounded here) show that the performance of the system (including entrainment limits) can be characterized with reasonable accuracy in terms of frequency *differences* alone. For these I have chosen $\omega_x - \omega_z$ and $\omega_y - \omega_z$. Furthermore, the mathematics show that changing the signs of both frequency differences (but preserving their magnitudes) leads to the same system performance (except for an inversion of the phase relationships with z). Consequently, the study for either one of the frequency differences can be restricted to only its positive or negative values. I have therefore chosen $\omega_y - \omega_z$ to have only negative values (the conventional situation), and $\omega_x - \omega_z$ is assigned both positive and negative values.

The results of the study for a weak Zeitgeber ($F_{zy} = 0.7$) are shown in Fig. 5. The contours in this entrainment diagram represent the values of frequency differences at which changes occur in the character of the system response to the imposed Zeitgeber. The areas of the diagram demarcated by the contours correspond to specific types of system responses and are designated by letter symbols. The meanings of these symbols are as follows:

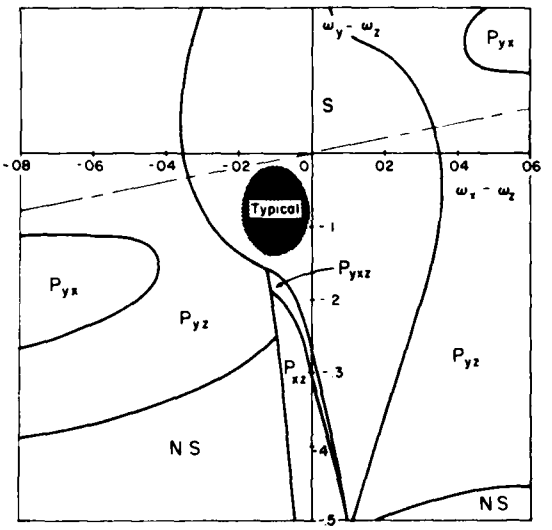


FIG. 5. The entrainment diagram for a system of two interacting Van der Pol oscillators with a sinusoidal Zeitgeber acting on *y* alone. The lettered symbols are explained in the text. Zeitgeber strength is 0.7.

S:	full synchrony; both x and y are entrained to z, and no phase modulations are seen
P_{yz}:	y is entrained to z, but x is not; because x has an average period different from that of z, the y variable shows phase-trapped modulations
P_{xz}:	x is entrained to z, but y is not; the x variable shows phase-trapped modulations
P_{xy}:	x and y have the same average period, which is different from that of z; both x and y show phase-trapped modulations
P_{xyz} or P_{yxz}:	x and y have the same average period as z, but they are only phase trapped to z, not synchronized; the first subscript denotes which of x or y has the stronger phase modulations
NS:	complete desynchronization; x, y, and z all have different average periods; both x and y show a complex mixture of the three periods

To facilitate the interpretation of Fig. 5, modified versions are shown as Figs. 6A and 6B. They emphasize, through shading, the range of frequency differences for which z entrains x or y, respectively. In both figures, the shaded area represents all those cases where the particular variable has the same average period as z. On all three diagrams, small ranges of ω_x and ω_y values are indicated that are presumed to be typical of healthy young adults functioning in the conventional 24-hr environment. Such subjects would, on release from Zeitgeber drive, be in an internally synchronized state, but close to the phase-trapped state.

The first thing to note in any of these figures is that the scale of $\omega_x - \omega_z$ is expanded five times with respect to the scale of $\omega_y - \omega_z$. This is because the effective isolation of x from z results in a much smaller range of entrainment for x than for y. As Fig. 6A shows, x can be entrained for $|\omega_x - \omega_z| < 0.036$ when $\omega_y = \omega_z$, and as $|\omega_y - \omega_z|$ increases, the range of entrainment progressively decreases. There is also a slight shift of the center of the entrainment band, so that for $\omega_y - \omega_z < 0.1$ the band is displaced about 0.005 in the direction of positive $\omega_x - \omega_z$. Qualitatively, the entrainment band for x is quite regular, and the results can be presented very simply. For ω_y in the "normal" range ($1 > \omega_y > 0.88$), x can be entrained by z, provided the intrinsic period of the x subsystem is within about ± 0.8 hr of the Zeitgeber period. For ω_y in the range appropriate to internal desynchronization ($\omega_y < 0.85$), the total range of period discrepancy between x and z for which entrainment occurs is no more than about 1 hr and is biased in favor of x periods that are shorter than z by about 10 to 15 min.

The entrainment of y is somewhat more complex. In crude terms, z can entrain y over a very broad frequency range $|\omega_z - \omega_y| \lesssim 0.4$. This magnitude is very close to what one would find if the x subsystem were to exert no drive onto y (i.e., $F_{xy} = 0$). However, the existence of the $x \rightarrow y$ drive produces some

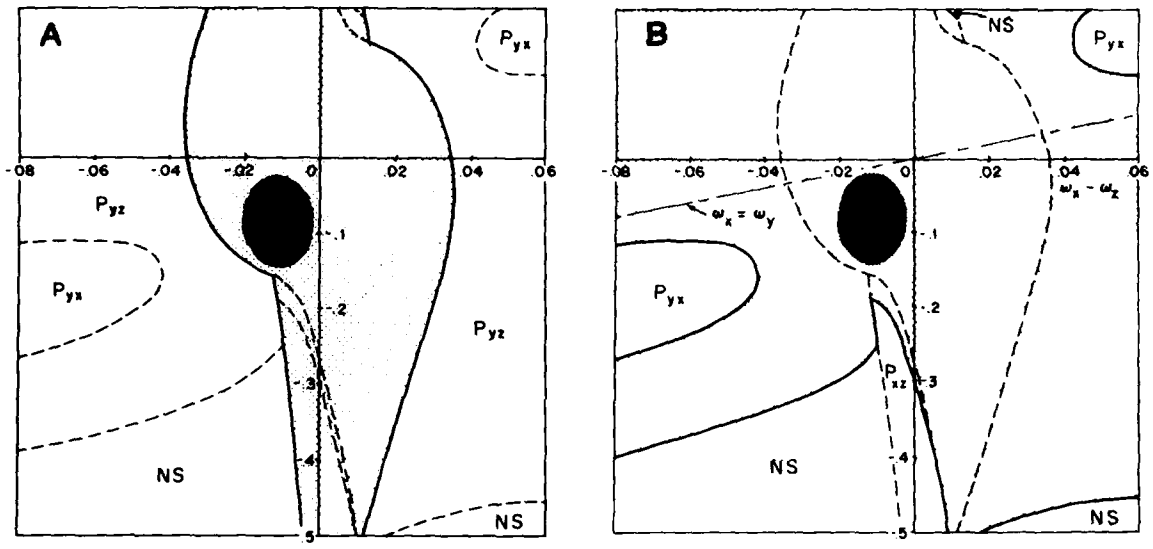


FIG. 6. A: Regions of *x* entrainment (*shaded*) taken from Fig. 5. B: Regions of *y* entrainment (*shaded*) taken from Fig. 5.

marked differences from such a simple picture. In the region where x is entrained by z and $\omega_x - \omega_z < 0$ there is a marked reduction in the ability of z to entrain y (this is represented by the tongue-shaped region labeled P_{xz}). This reduction is due to opposition between the $x \rightarrow y$ drive and the direct $z \rightarrow y$ drive. The other marked effect of the $x \rightarrow y$ drive is the creation of the P_{yx} regions. In these regions, y prefers to be entrained to x (i.e., phase trapped to x) rather than entrained to z. Interestingly, these regions do *not* correspond to the condition $\omega_x = \omega_y$, as might have been supposed, but are displaced from this locus in the direction of larger $|\omega_y - \omega_z|$. In fact, the existence of these P_{xy} regions must be counted one of the least easily explained phenomena of this not-too-complicated system. What is particularly surprising is that at any fixed $\omega_x - \omega_z < 0.042$, as $\omega_y - \omega_z$ is progressively decreased below zero, the y subsystem prefers first to be linked with z, then to be linked with x, and then again linked with z before, finally, desynchrony occurs. Because the progressive detuning between y and z (i.e., increase in $|\omega_y - \omega_z|$) is accompanied by an *equal* progressive detuning between y and x (increase in $|\omega_y - \omega_x|$), it is not at all obvious why the y entrainment preference should alternate in this way.

Probably the most remarkable result displayed in this entrainment diagram is the existence of the extensive P_{xz} regime, where the failure of y to be entrained occurs for values of $\omega_y - \omega_z$ that are well within entrainment limits if the $x \rightarrow y$ drive is absent. The x subsystem is actively suppressing y entrainment, while at the same time x itself is entrained. What is also very interesting is that this remarkable situation occurs for values of the frequency differences that are not far from "typical" values. Suppose the "typical" subject were to be released into free run and his ω_y were then to decrease secularly, so as to bring about internal desynchrony. Such a subject would then have exactly the proper ω_x and ω_y values to exhibit P_{xz} when a relatively weak 24-hr light/dark cycle is reinstated.

Between the P_{xz} regime and the S regime the entrainment diagram shows a narrow band of P_{yxz}. There is an analogy between this and the free-run situation where the phenomenon of phase trapping is interposed between synchrony and desynchrony. Here, as y is breaking away from the xz alliance, there is a band in which y is phase-trapped to xz before the complete break occurs.

It is instructive to see the actual forms that the computer simulations take for various cases of partial entrainment. Figures 7A and 7B show examples of P_{yz} and P_{yx}, respectively. In both of these simulations the period of the Zeitgeber was maintained at 24 hr ($\omega_z = 1$), and the appropriate frequency differences, $\omega_x - \omega_z$ and $\omega_y - \omega_z$, were achieved by adjusting ω_x and ω_y. Consequently, ω_x is below what I would consider the normal physiological range (which is a very small range). The simulations could alternatively be taken to represent a normal subject exposed to a Zeitgeber of period less than 24 hr (approximately 23 hr), for which the plotting has been made with respect to the 23-hr reference. The Zeitgeber is a cosine function of amplitude 0.7, with its positive maximum occurring in the center of each time band. Thus, it is convenient to think of the

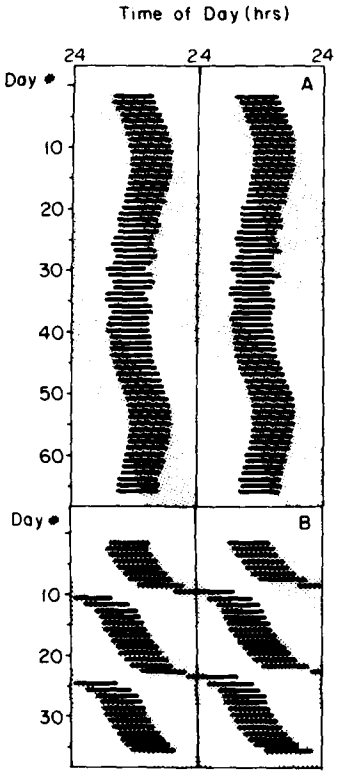


FIG. 7. A: A simulation demonstrating P_{yz}. For this example, $\omega_z = 1.0$, $\omega_x = 0.98$, and $\omega_y = 0.77$. The Zeitgeber is sinusoidal, with strength 0.7, and acts on y only. The center of each band corresponds to the time when z has a positive maximum. Plotting convention as in Fig. 4B. B: A simulation demonstrating P_{yx}. For this example, $\omega_z = 1.0$, $\omega_x = 0.94$, and $\omega_y = 0.85$. Except for ω_x and ω_y, everything is as in A. Plotting convention as in Fig. 4B.

center of each band as representing "dawn," or the principal awakening stimulus.

Consider P_{yz} first (Fig. 7A). Here x displays its intrinsic period (which is longer than the period of z), although there are significant modulations of its phase drift. The regions of x below the mean tend to drift more slowly at the phase position in which they occur shortly after "dawn." Sleep generally occurs in approximately the normal relation to the "dawn" stimulus. It can, however, be drawn off to later daily times by the influence of x, when x reaches a phase relative to sleep that is close to that normally observed in synchrony. It is especially interesting to note that when the Zeitgeber and x work in concert on y (as for days 0 to 15, for example), the sleep episodes occur with a smooth regularity. However, when their effects oppose (as for days 16 to 38), the Zeitgeber is able to enforce only a phase-trapping constraint on y. This phase-trapping cycle has a period of 4.5 days.

In the example of P_{yx} (Fig. 7B) there are some evident effects of z on x, but much less than those seen in P_{yz}, because y is itself no longer following z.

Although y is following x (on average), the effect of z on y is very strong, and there is the suggestion that y is periodically almost "captured" by z. This occurs, understandably, when x has drifted to the phase position with respect to z that it assumes for normal entrainment to a 24-hr Zeitgeber. It is at this phase that x and z act on y in a cooperative way.

Figure 8 shows an example of a system with no mutual entrainment whatsoever (*NS*). For this simulation it is convenient to think of the Zeitgeber as having the normal 24-hr period and the intrinsic period of x as being within the normal range (24.5 hr). However, the intrinsic period of y is much longer than normal (40 hr). This period is one that is seen in some internally desynchronized free-running subjects, usually after the desynchrony has persisted for a long time. Thus, the simulation can be thought of as what might be found if such a free-running subject were to be exposed to a relatively weak 24-hr Zeitgeber. As in Figs. 7A and 7B, it is clear that when the drift of x brings its phase, relative to z, to the value it has under ordinary entrained conditions, the action of x together with z makes y appear to be phase-trapped to z (with a modulation cycle of about 3 days). However, as x drifts away from this special

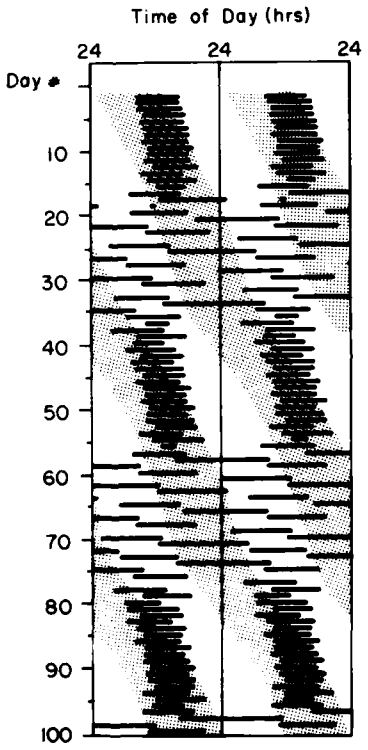


FIG. 8. A simulation demonstrating *NS*. For this example, $\omega_z = 1.0$, $\omega_x = 0.98$, and $\omega_y = 0.63$. All other properties and conventions as in Fig. 7.

phase, y breaks loose and exhibits its own intrinsic period. Thus, the *NS* state of Fig. 8 is readily seen to be similar to P_{yz} of Fig. 7A, except that the ability of z and x to influence y is weaker for Fig. 8 by virtue of the greater disparity of the y and z periods.

To demonstrate the rather unusual system behavior of P_{xz}, a simulation is chosen (Fig. 9) in which the z period is 24 hr and the intrinsic x period is 24.24 hr (a normal value), but the intrinsic period of y is allowed slowly to lengthen, increasing from 28.24 hr to 30 hr over 100 days. At the outset, both x and y are entrained to z in a very conventional set of relative phases. As the y period lengthens, the phases of both y and x are delayed progressively with respect to "dawn," until at day 23 the model subject does not awaken until about 6 hr after "dawn." But with further increase in the y period, the phase of y advances, while the phase of x is delayed further. Then, because the effect of x on y now tends to oppose the effect of z on y, y begins to break loose from the combined influences. First, y falls into phase trapping with respect to the x,y group (P_{yxz}), and as the phase-trapping amplitude grows, it ultimately develops into a loss of entrainment for y, (P_{xz}). This sequence bears a very close resemblance to the sequence of synchrony, phase trapping, and desynchrony seen in free-running subjects as the period of y is lengthened (Fig. 4B). Indeed, the loss of entrainment takes place at $\hat{\tau}_y = 29.3$ hr, which is closely comparable to the value at which desynchrony is often seen in free run. There is, however, a very important difference from the free-run situation because of the phase of x. Mid-low x occurs *after* sleep in the P_{yxz} regime, whereas it occurs *before* sleep in the free-run phase trapping. The reason for this reversal, according to the model, is that it is z that exercises the dominant influence over y in both P_{yxz} and P_{xz}, whereas in free run there is no z, and y is necessarily under the influence of x.

Figure 9 illustrates that in P_{xz} the sleep episodes cluster about the middle of each 24-hr interval, which is the time of Zeitgeber "dawn" (peak z drive). The longest sleep episodes are centered about 12 hr after this. Also, there is a band of 6-hr width that begins about 14 hr after "dawn" and ends about 4 hr before "dawn," during which no awakening takes place. It is this suppression of awakening that gives rise to the long sleep episodes. It is very interesting to make a comparison of this phenomenon with the human example of P_{xz} given by Wever (ref. 6, Fig. 81, and shown earlier here as Fig. 3). These data suggest that the subject was never entrained to the Zeitgeber and became internally desynchronized at day 11, at which point body temperature became entrained to z. These data also show that there is a time band lying between the hours 1900 and 2400 during which no awakening is observed. If "dawn" is taken to be the time when the lights of the Zeitgeber are turned on, this corresponds to hour 0400 in the experiment. The no-awakening band is then seen to extend from 15 hr after "dawn" to 4 hr before "dawn," in very good agreement with the model. The temperature data reported in Fig. 3 give only the times at which maxima and minima were observed, and it is very unreliable to ascertain cycle

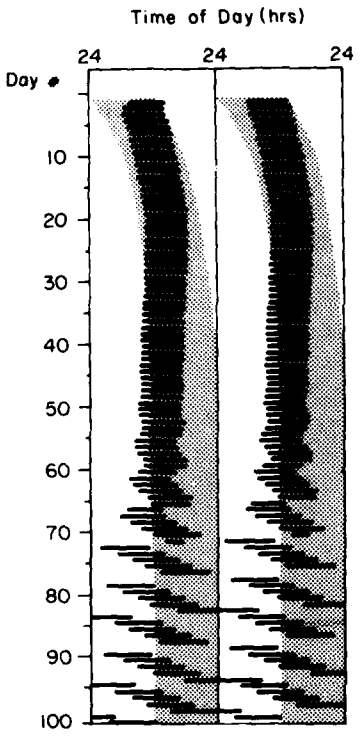


FIG. 9. A simulation demonstrating the transition sequence: S, P_{yxz}, P_{xz}. For this example, $\omega_z = 1.0$, $\omega_x = 0.99$, and ω_z decreases uniformly from 0.85 to 0.80. All other properties and conventions as in Fig. 7.

phase from such scanty evidence. Bearing in mind this caveat, low temperature occurs close to "dawn," whereas the model places mid-low x some 6 hr later. Partial explanation of the discrepancy may lie in the sleep-evoked component of T_c. This component has a minimum of evoked T_c at the time of sleep onset. Because the clustering of sleep episodes occurs centered at "dawn," the sleep-evoked T_c minimum will generally occur some 4 hr or so ahead of "dawn," and consequently the minimum of T_c, accounting for the effects of both x and sleep, will be well in advance of mid-low x.

Two other sets of human data taken from the work of Wever (ref. 6, Fig. 71, 69) are shown as Figs. 10A and 10B. In Fig. 10A, when the Zeitgeber period is made short (22.67 hr), the subject's activity cycle follows the imposed light/dark cycle, but T_c is no longer entrained and instead follows a 24.8-hr period. This is P_{yz}. In Fig. 10B, for the same 22.67-hr Zeitgeber, the subject's activity and T_c follow a 25.2-hr average cycle length. This is P_{yx}. Because we have no independent calibration of what strength the Zeitgeber of these experiments should represent in terms of the model system, comparison between these results and model results can be only semiquantitative. However, note that $\tau_z = 22.67$ corresponds to $\omega_z = 1.059$, and for either data set, $t_x = 24.8$

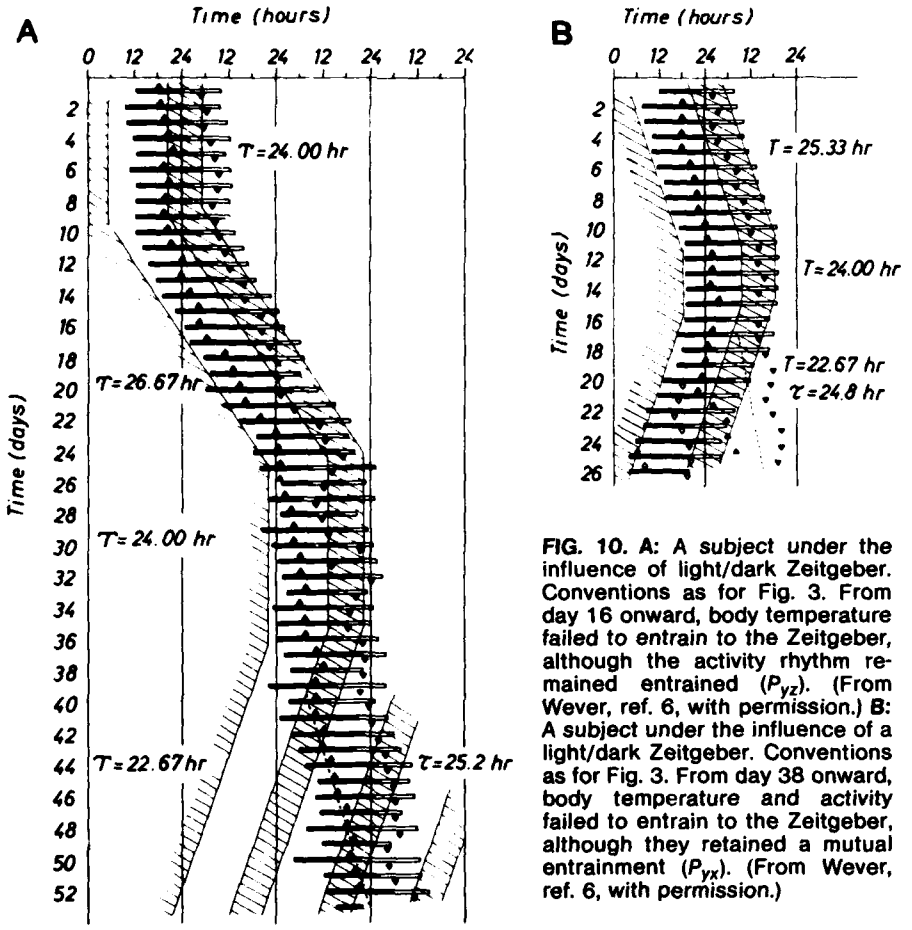


FIG. 10. A: A subject under the influence of light/dark Zeitgeber. Conventions as for Fig. 3. From day 16 onward, body temperature failed to entrain to the Zeitgeber, although the activity rhythm remained entrained (P_{yz}). (From Wever, ref. 6, with permission.) B: A subject under the influence of a light/dark Zeitgeber. Conventions as for Fig. 3. From day 38 onward, body temperature and activity failed to entrain to the Zeitgeber, although they retained a mutual entrainment (P_{yx}). (From Wever, ref. 6, with permission.)

($\omega_x \cong 0.97$) is a reasonable estimate. This gives $\omega_x - \omega_z = -0.09$, which lies a bit off the left limit of the entrainment diagram, Fig. 5. It is clear that either P_{yz} or P_{yx} can be found at this value of $\omega_x - \omega_z$ for values of ω_y within the normal range. A value of $F_{zy} = 0.7$ was used in the simulations from which Fig. 5 was deduced. In very crude terms, increasing F_{zy} by some small multiple of 0.7 will expand the scale of the response diagram by that same multiple. Therefore, F_{zy} can be at least doubled without losing the two alternative responses, P_{yz} or P_{yx}, at that value of $\omega_x - \omega_z$. In short, the fact that two different subjects displayed the two different responses is attributable to different intrinsic ω_y. But either of those ω_y values is within what may be regarded as the normal range. Furthermore, this explanation is not dependent on choosing a particular F_{zy} in the model but is valid over a considerable range of Zeitgeber strengths.

SUMMARY AND PROJECTIONS

From a large body of human free-run data it is possible to extract general features that any successful model of the human circadian system must reproduce. Once the existence of two rhythmic subsystems is acknowledged, probably the most significant feature of the data in terms of modeling implications is the phenomenon of phase trapping, which actually represents a limited loss of synchrony between the subsystems.

A very simple model consisting of two quasilinear oscillators, interacting mutually through linear coupling, is able to mimic the requisite general features provided that one parameter (the intrinsic period of the y oscillator) is assigned a secular variation. The physiological basis of the drifting y period is unknown. If this simple model system is subjected to a periodic excitation via the y subsystem, a varied collection of full and partial entrainment responses is elicited; these depend on the differences between the intrinsic periods of the subsystems and the period of z. These model responses have their counterparts in various human responses seen in experiments with light/dark Zeitgebers, thereby lending an encouraging qualitative (and semiquantitative) support to the basic model concept.

Further progress in modeling appears to require two major experimental thrusts. The first is an effort to quantify Zeitgeber strength through a careful series of experiments designed to determine entrainment limits. If the model is assumed to be fundamentally correct as it now stands, simulations show that internal adjustments of the system take place very slowly, and therefore they imply that experimental determination of entrainment limits is liable to serious error unless special precautions are taken. It therefore seems reasonable to use model simulations to guide the design of these difficult experiments.

The second experimental development should be an effort to obtain a better representation of x than T_c alone affords. One approach is to assume that T_c represents a simple sum of the x and y influences. If, further, the y influence is presumed to be timed by the partitioning of the activity cycle into sleep and wake, but otherwise independent of the magnitude of y, the influence can be determined by correlation methods, using the timing of sleep to provide markers of cycle phase. Then the y-evoked contribution to T_c can be subtracted. The remainder is expected to be a better representation of x than T_c itself, and comparison with model simulations can be used to test this hypothesis.

REFERENCES

1. Aschoff, J., and Wever, R. (1976): Human circadian rhythms: A multioscillatory system. *Fed. Proc.*, 35:2326–2332.
2. D'Heedene, R. N. (1959): On limit sets of n-th order ordinary differential equations. Ph.D. thesis, Division of Engineering and Applied Physics, Harvard University.

SUMMARY AND PROJECTIONS

From a large body of human free-run data it is possible to extract general features that any successful model of the human circadian system must reproduce. Once the existence of two rhythmic subsystems is acknowledged, probably the most significant feature of the data in terms of modeling implications is the phenomenon of phase trapping, which actually represents a limited loss of synchrony between the subsystems.

A very simple model consisting of two quasilinear oscillators, interacting mutually through linear coupling, is able to mimic the requisite general features provided that one parameter (the intrinsic period of the y oscillator) is assigned a secular variation. The physiological basis of the drifting y period is unknown. If this simple model system is subjected to a periodic excitation via the y subsystem, a varied collection of full and partial entrainment responses is elicited; these depend on the differences between the intrinsic periods of the subsystems and the period of z. These model responses have their counterparts in various human responses seen in experiments with light/dark Zeitgebers, thereby lending an encouraging qualitative (and semiquantitative) support to the basic model concept.

Further progress in modeling appears to require two major experimental thrusts. The first is an effort to quantify Zeitgeber strength through a careful series of experiments designed to determine entrainment limits. If the model is assumed to be fundamentally correct as it now stands, simulations show that internal adjustments of the system take place very slowly, and therefore they imply that experimental determination of entrainment limits is liable to serious error unless special precautions are taken. It therefore seems reasonable to use model simulations to guide the design of these difficult experiments.

The second experimental development should be an effort to obtain a better representation of x than T_c alone affords. One approach is to assume that T_c represents a simple sum of the x and y influences. If, further, the y influence is presumed to be timed by the partitioning of the activity cycle into sleep and wake, but otherwise independent of the magnitude of y, the influence can be determined by correlation methods, using the timing of sleep to provide markers of cycle phase. Then the y-evoked contribution to T_c can be subtracted. The remainder is expected to be a better representation of x than T_c itself, and comparison with model simulations can be used to test this hypothesis.

REFERENCES

1. Aschoff, J., and Wever, R. (1976): Human circadian rhythms: A multioscillatory system. *Fed. Proc.*, 35:2326–2332.
2. D'Heedene, R. N. (1959): On limit sets of n-th order ordinary differential equations. Ph.D. thesis, Division of Engineering and Applied Physics, Harvard University.

3. Fuller, C. A., Lydic, R., Sulzman, F. M., Albers, H. E., Tepper, B., and Moore-Ede, M. C. (1981): Circadian rhythm of body temperature persists after suprachiasmatic lesions in the squirrel monkey. *Am. J. Physiol.*, 241:R385–R391.
4. Iouye, S. T., and Kawamura, H. (1979): Persistence of circadian rhythmicity in mammalian hypothalamic island containing the suprachiasmatic nucleus. *Proc. Natl. Acad. Sci. USA*, 76:5961–5966.
5. Kronauer, R. E., Czeisler, C. A., Pilato, S. F., Moore-Ede, M. C., and Weitzman, E. D. (1982): Mathematical model of the human circadian system with two interacting oscillators. *Am. J. Physiol.*, 242:R3–R17.
6. Wever, R. A. (1979): *The Circadian System of Man*. Springer-Verlag, New York.
7. Czeisler, C. A. (1978): Human circadian physiology: Internal organization of temperature, sleep–wake and neuroendocrine rhythms monitored in an environment free of time-cues. Ph.D. thesis, Department of Neuro- and Biobehavioral Sciences, Stanford University.
8. Moore, R. Y., and Leen, N. J. (1972): A retinohypothalamic projection in the rat. *J. Comp. Neurol.*, 146:1–14.
9. Klein, K. E., Wegemann, H. M., and Hunt, B. I. (1972): Desynchronization of body temperature and performance circadian rhythms as a result of outgoing and homegoing transmeridian flights. *Aerospace Med.*, 43:119–132.
10. Kronauer, R. E., Czeisler, C. A., Pilato, S. F., Moore-Ede, M. C., and Weitzman, E. D. (1983): Mathematical representation of the human circadian system: Two interacting oscillators which affect sleep. In: *Sleep Disorders: Basic and Clinical Research*, edited by M. H. Chase and E. D. Weitzman, pp. 173–194. Spectrum, New York.

DISCUSSION

Dr. Pastel: In all of these circadian studies the subjects are told not to sleep unless they are going to sleep for a long period of time—in other words, they are asked not to nap. However, many people have a tendency to nap in the afternoon. Would your model account for that?

Dr. Kronauer: No. We have not got such a mechanism in our model. But I think there is a little misconception. The subject can go to sleep whenever he wishes. However, there is a certain ritual involved. He has to give notice, be fitted with the electrodes, and change into his pajamas. As a result, there is less tendency to go to sleep for casual reasons.

Dr. Wever: Our model has only one input, simultaneously for phasic and tonic effects, and it can model very nicely *all* effects. So, why complicate the system? There is another way my model is different. My impedance ratio is 1, but there is a difference in the amplitudes.

Dr. Kronauer: There are two different ways of scaling. I happen, arbitrarily, to have chosen the scales where the oscillations both have an amplitude of 2, i.e., the nominal amplitude of my oscillators is 2. Consequently, I put the difference in the impedance ratio. Alternatively, I could have put it in as a change in amplitude and it would have been the same thing.

Dr. Wever: No. It is not the same. It is the same in all the experiments you have shown. But there are many additional experiments which you cannot explain with an impedance ratio of 12 or 14 or 4, but they must have the same amplitude.

Dr. Kronauer: Excuse me, but with the Van der Pol oscillator they are identical.

Dr. Wever: Yes, with the Van der Pol, but that is one reason why I think the Van der Pol oscillator is not a good model. May I ask another question? You showed very nicely that it is much better to have the Zeitgeber influence only the y oscillator. Why have you not tried to influence both? Then you get a much better coincidence.

Dr. Kronauer: Let me respond. Don't misunderstand me. I'm not trying to say that it is not possible, through one agency or another, to have an influence on the x oscillator directly—chemicals could perhaps do it.

Dr. Wever: But even the subjects show that.

Dr. Kronauer: The results that I've seen of your work are in fact the only extensive results of Zeitgeber effects that I know of, and they appear to me to be totally consistent with a drive solely on y. That is a simple statement, and I think we have much more work to do with various timekeepers before we can answer the question fully.

Dr. Wever: As you know, we have thought for a long time in the same way as you. But other people have suggested that different Zeitgebers influence the two oscillators differently. I have the feeling that the simplest way that is consistent with all the different Zeitgeber modes of action, whether they be social cues or light/dark cycles, is that the Zeitgebers influence both oscillators, and to the same degree. That gets the best coincidence with all experiments of very different types.

Dr. Kronauer: If we could accept other features—for example, that the way the Van der Pol oscillators couple represents the real internal structure of the system—then we would be in the position to design specific experiments as to whether Zeitgeber A or Zeitgeber B affects x or affects y. That is what I would hope would be one of the results of utilizing a model of this kind.

Dr. Wever: Have you superimposed random noise in your modeling?

Dr. Kronauer: No, but the system is robust. That is to say, both of these oscillators are very stable and will come back to their mean values. So the addition of noise will not produce any untoward, remarkable changes.

Dr. Lerman: Can noise send the oscillators outside their own boundaries?

Dr. Kronauer: Oh yes. These are all predictions from a deterministic model. And, in fact, I might comment that it is extremely difficult to identify those boundaries. As Dr. Gander will be glad to tell you, if you are trying to drift across one of those boundaries, one must change the parameters *extremely* slowly. And it turns out that if that is the way the human system works, to find a desynchronization boundary you have to run experiments that change the Zeitgeber less than a minute a day, or else you will not get the correct boundary.

Dr. Enright: If I understood correctly, the feedback of the temperature oscillator on the sleep oscillator is through the rate of change of temperature, rather than the temperature itself. To me, biologically it seems more plausible that absolute body temperature would affect the time to go to sleep, much more than the rate of changes.

Dr. Kronauer: I agree. There is a degree of arbitrariness in the way you construct the coupling. But you have to be careful. If you do not assign the correct signs of those couplings, you will find that the system will not come to a compromise. So, in fact, there's a certain sign ratio that must be observed. Second, it would work perfectly well if I had x in the one and y in the other. The only thing is, the absolute phases would then be shifted about 90°. This particular choice of variables gave what were the correct phases.

Mathematical Models of the Circadian Sleep-Wake Cycle, edited by M. C. Moore-Ede and C. A. Czeisler. Raven Press, New York © 1984.

Circadian Gating of Human Sleep–Wake Cycles

*Serge Daan and **Domien Beersma

**Zoological Laboratory and **Department of Biological Psychiatry, University of Groningen, The Netherlands*

CIRCADIAN MODELING: AIMS AND STRATEGY

The temporal organization of behavior in animals, including humans, presents one of the major challenges today for both physiological and functional analysis. Circadian processes play a key role by their function in integrating behavioral organization into the predictable time patterns of the environment on a rotating planet. Two aspects of circadian rhythmicity make its analysis exceedingly complicated: first, the fact that such rhythms behave as self-sustained, nonlinear oscillations (40); second, the recognition that they may be the combined output of several structures (probably a multitude of structures) with such oscillatory capacity, in continuous interaction with each other and with environmental periodicities. This complexity has called for extensive mathematical modeling of the circadian system (19,22,32,38,39,51,53,55) to help in understanding the relationships among empirical results and occasionally to suggest new experimental designs. Of the many models developed, few have been refuted by subsequent experimental analysis. Most models still survive in peaceful coexistence, albeit with fluctuating popularity. Indeed, the mainstream of current physiological research in circadian rhythms is hardly influenced by mathematical simulations. In our view, this is because of a superabundance of parameters in many of the models and because of the absence of specific hypotheses regarding the physiological equivalents of those parameters [with the notable exception of Enright's model (22)]. The physiologist hoping to gain further insight into a system is left without guidance from the mathematicians as to how to test their models. There is a continuing danger of overcomplication and overabstraction in mathematical modeling. A useful model should be minimally complex to account for an existing set of data and maximally specific about what its parameters mean in physiological terms. It should not aim at completeness. The essence of a model's usefulness is in being a simplification of nature, rather than in approaching the complexity of nature itself.

The aim of this chapter is to propose a simple hypothesis of the generation of human sleep–wake rhythms and to follow its complex consequences with the

help of mathematical simulations. There are several precursors of this model, and their essential characteristics can be summarized as follows:

Enright (22) designed an elegant model of a circadian pacemaker, composed of a multitude of neuronal elements, each capable of rather imprecise circadian firing oscillations, and together forming an ensemble with very precise self-sustained circadian properties, matching in great detail what is known of vertebrate activity rhythms. Enright's assumption is that the activity of such a pacemaker will directly elicit activity of the organism, whereas rest of the pacemaker will lead to rest of its bearer.

Wever (53), in addition to extensively reviewing the main body of data obtained in 20 years of isolation experiments, proposed a general mechanism for the generation of human circadian rhythms. In this concept, there are at least two self-sustained oscillators involved, one stronger than the other, and mutually interacting in such a way that they are normally running in synchrony. The strong oscillator controls rhythms in body temperature; the weak oscillator controls rhythms in sleep and wakefulness. The latter oscillator, although normally synchronized to the frequency of the stronger one, may in free-running conditions occasionally dissociate from the latter and exhibit its own frequency.

Kronauer et al. (32) presented simulations based on a mathematical formulation of this model using two coupled Van der Pol oscillators. Using six variable parameters, these authors were able to select values such that a reasonable approximation of observed sleep–wake rhythms was obtained.

Eastman (21; and *this volume*) stated that the assumption of a second oscillator is not required to explain these observed patterns. In her view, occasional spontaneous extensions of the activity time, together with a strong feedback effect of sleep and wakefulness on the single basic oscillator, may be sufficient assumptions to account for the data. If such extensions occur regularly, however, such as in "circabidian" rhythms, Eastman's model also requires a second oscillator (the "phase-shifting oscillator").

Of these models, Enright's is the only one that makes specific assumptions about the (neuronal) structure generating sleep and wakefulness. The other models are abstract, and their specific parameters, insofar as defined, have no obvious physiological counterparts. Enright's model, designed as a general hypothesis of circadian sleep–wake pacemakers, is also the only one at variance with the human data, unless an additional oscillator, controlling body temperature and interacting with the sleep–wake pacemaker, is invoked. None of the models considers the body of knowledge on sleep per se, its temporal structure, and experimental manipulation of sleep. The simple fact that sleep in humans is not fixed exclusively by deterministic oscillatory structures, but may to a large extent be modified consciously, is also rarely considered (however, see C. Eastman, *this volume*).

Dissatisfied with this situation, we have attempted to formulate a minimally complex hypothesis incorporating known homeostatic properties of sleep with a

circadian regulatory mechanism. The model formulated and used for simulations to study its behavior invokes a single circadian oscillator or pacemaker, as suggested by Eastman (21; and *this volume*). In addition, it includes a homeostatic regulatory process of sleep and wakefulness with empirically estimated parameters along the lines proposed by Borbély (9). This model, essentially using two free parameters, is able to match most of the available data. We see no evidence in the literature at variance with the model, but we hope it will help to generate such evidence in the future.

SLEEP: HOMEOSTATIC VERSUS CIRCADIAN PROCESS

Before elaborating on the construction of our model, a digression on the general nature of rhythms in sleep and wakefulness is appropriate. In order to truly understand such rhythms, we need to examine the contribution of sleep and wakefulness to evolutionary fitness, the final common path in biological analysis. However, sleep, more than any other element in animal behavior, has eluded functional understanding. As recently as 1979, in the introduction to a symposium on the functions of sleep, Rechtschaffen stated that "we do not know why so much of our own lives, the lives of all the mammals and, very likely submammalian species as well, should be captured by sleep" (44). Yet speculations and hypotheses have been manifold. In a useful review, Webb (48) arranged the existing theories into five categories, admittedly with some measure of overlap. Without attempting to rehearse them in great detail, we summarize them as follows:

Restorative theories. Most staunchly defended by Hartmann (28), such theories presume that sleep restores physiological properties of some body tissues (often restricted to the central nervous system or parts of it) and that such restoration is needed for their proper functioning during the active phase. Specific hypotheses are rarely articulated.

Protective theories. Along similar lines, Pavlov [cited in Webb (48)] argued that sleep suppresses cortical activity and thereby prevents the exhaustion of cortical neurons.

Instinctive theories. In this view, taken, for instance, by Moruzzi [cited in Webb (48)], sleep is seen as a consummatory behavior associated with an innate organized pattern.

Energy-conservation theories. Sleep researchers such as Berger (7) have stressed the possible energy savings that warm-blooded animals achieve during sleep by turning their thermostats down and remaining inactive.

Immobilization theories. Recent ideas expressed by Webb (47) and Meddis (33,34) presume that sleep enhances survival by rendering animals immobile and nonresponsive, thereby increasing their safety, during ecologically adverse parts of the day–night cycle.

These approaches are obviously not mutually exclusive; in fact, the instinct theories represent a way of describing sleep behavior rather than defining

benefits to the organism arising from it. The extreme positions stress the possible benefits with respect to either the internal milieu or the external periodic environment. The more classic views (i.e., restorative theories, protective theories) of sleep as a process of recovery from the fatigue created by prior activity are associated with the long-term maintenance of some homeostatic balance within the organism. We shall call these the homeostatic theories. In the energy-conservation and risk-reduction theories, the timing of sleep with respect to the day–night cycle is crucially important, and hence emphasis is placed on the circadian aspect of function.

Sleep-deprivation experiments have played a significant role in the controversy between the homeostatic and circadian approaches. The increased tendency to sleep and the "rebound" observed in EEG analysis after a night of sleep deprivation present a problem for the circadian theories, because the behavior should be optimized with respect to time of day regardless of the animal's prior history. On the other hand, sleep deprivation is never completely compensated for by subsequent sleep, as would be predicted by the homeostatic theories. The general negative correlation between activity time and subsequent sleep duration (53) would similarly not be expected on the basis of a homeostatic recovery process. However, none of these arguments completely excludes either type of function. That sleep duration is not positively associated with prior wake time, natural or forced, is easily accommodated if the restorative processes are nonlinear. The rebound after sleep deprivation may likewise be explained as resulting from the motivational mechanism controlling sleep without contradicting its circadian function, as Meddis (33) has persuasively argued.

The proponents of the homeostatic and circadian theories of sleep function certainly agree in one respect: The pressure to sleep builds up during activity and during enforced sleep deprivation. The difference is that this is seen either as crucial for sleep function or merely as part of the sleep-generating mechanism (Fig. 1). Although the circadian control of sleep and wakefulness is beyond doubt, recent evidence suggests that a homeostatic rebound after sleep deprivation is still seen in rats with lesions of the suprachiasmatic nuclei that have destroyed any circadian rhythmicity, including that of sleep and wakefulness (12,35).

The contributions of sleep to survival may eventually be evaluated by manipulating sleep behavior in animals in their natural environment and studying the consequences for the number of offspring propagated. But until this formidable task is undertaken, we do well to accept the available evidence for both the circadian control and the restorative aspects of sleep. In an integrated view, we can acknowledge that the periodic reduction of energy expenditure associated with lowered metabolic rate and inactivity at times of day when activity would be inefficient may benefit the animal. In some cases, immobility may further reduce predatory risks, although one would not expect reduced sensory perception at times when the dangers of predation are maximal. The

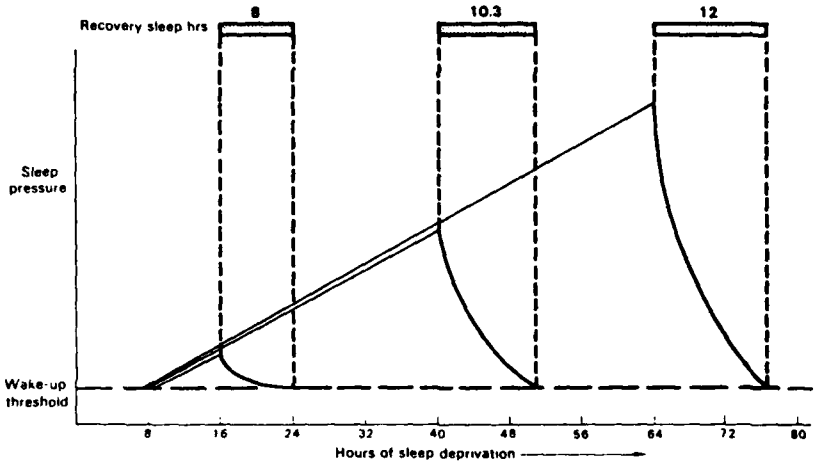


FIG. 1. Meddis's explanation of incomplete compensation of sleep deprivation. The increase in sleep time may not be proportional to sleep deprivation if the rise in sleep pressure (during waking) is linear and the decrease (during sleep) is exponential. (From Meddis, ref. 34, with permission.)

reduction of energetic expenditure during sleep may have enabled animals to increase their bodily activity and central nervous system activity in the active part of the circadian cycle above the maximum level that could be sustained on a permanent basis had sleep not occurred. In this sense, sleep may now serve to restore a homeostatic balance, although the nature of the physiological deterioration in the absence of sleep remains virtually unknown. Even if the original contributions of sleep have been to ecological efficiency, this itself may possibly have permitted overstraining the organism's tissues during activity to a level where subsequent sleep is normally indispensable.

Thus, although there is reason to take an integrative view on the *function* of sleep, careful studies by Borbély (8–10) have recently been leading to a synthesis of homeostatic and circadian aspects in the *mechanism* generating sleep and wakefulness. The model proposed by Borbély on the basis of his sleep-deprivation experiments is essentially one of a circadian oscillator interacting with a self-regulated homeostatic process (Fig. 2). Borbély proposed that there is a sleep-regulating variable that increases in strength during wakefulness and decreases exponentially during sleep. Onset and cessation of sleep are largely determined by a threshold that oscillates in circadian fashion (Fig. 2). Borbély did not specify the variable, but suggested that a humoral CSF factor (11,23,36,37) with a concentration dependent on prior waking time is associated with it. It is further reminiscent of the "sleep pressure" in Meddis's model (Fig. 1), although there the threshold for sleep onset was not identified.

Our model for the human system relies heavily both on Borbély's data and on

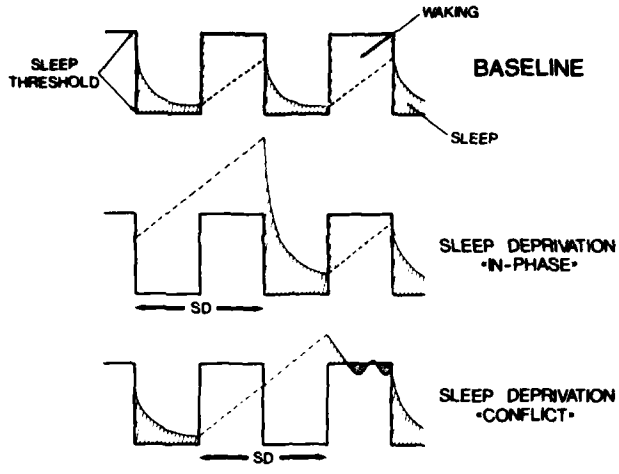


FIG. 2. Borbely's model of sleep regulation based on the combination of a circadian oscillator and a self-regulated process. A single threshold varies with circadian frequency. A rise of the sleep-regulating variable above this threshold triggers sleep, a (exponential) drop below the threshold terminates sleep. (From Borbely, ref. 9, with permission.)

his general integrative views of sleep and wakefulness. For present purposes, it makes no difference if the breakdown of the "sleep-regulating variable" itself confers the ultimate survival value of sleep behavior or, alternatively, if it is part of the mechanism ultimately leading to energy savings and reduction of natural risks. Whatever the ultimate value, it can be conferred to the organism only if the entire process is optimally timed in the external day–night cycle. There is precedence for the daily timing of spontaneous processes in the pupal eclosion of *Drosophila*. In an elegant series of experiments, Pittendrigh and Skopik (42) have shown that among the various developmental processes during larval and pupal life, there is one that is "gated" by a circadian oscillator, running in synchrony with the light–dark cycle and guaranteeing that adult flies emerge during a 6-hr "gate" opened around sunrise. Similarly, recordings hamster activity presented by Davis and Menaker (20) strongly suggest that a gate is opened once during every circadian cycle for the expression of higher-frequency periodicities in locomotor activity. Our model assumes that sleep and wakefulness are essentially homeostatically self-regulated processes varying between upper and lower thresholds, but that a circadian oscillation in these thresholds gates sleep such that it normally coincides with the night. We shall make no assumptions about the circadian pacemaker that regulates this gating, how it is composed of different elements, or how it is synchronized by light and darkness. It may be essentially a multiunit oscillator of the type proposed by Enright (22). Our only concern here will be with its output on the thresholds for

the sleep-regulating variable, how it coordinates the homeostatic process of sleep, and how it may under some circumstances lose control.

A HUMAN CIRCADIAN GATING MODEL

Our first assumption is that sleep is regulated by a variable S (Borbély's sleep-regulating variable) that builds up during activity and breaks down during sleep. Following Borbély (9), we have further assumed that the breakdown of S is reflected in characteristics of the sleep EEG. The kinetics of the buildup and breakdown processes can then be derived from the data of Borbély et al. (10) on the temporal course of the integrated EEG power density during sleep in humans, as affected by sleep deprivation (Fig. 3).

The decrease in power density in the 0.7- to 25-Hz domain during sleep, primarily due to a decrease in slow-wave sleep (0.7–2.5 Hz), is essentially exponential. Thus, we have simulated the S breakdown process by the exponential equation

$$S_i = (d)^i \cdot S_e \tag{1}$$

where S_i is the value of S at i time units after the onset of sleep, and S_e is the value of S at sleep onset (which equals end of activity). We use the half hour as the time unit for the process and derive for the breakdown rate d a value of 0.888 from the slope of Borbély's oblique line (−0.0515 log units/half hour) in

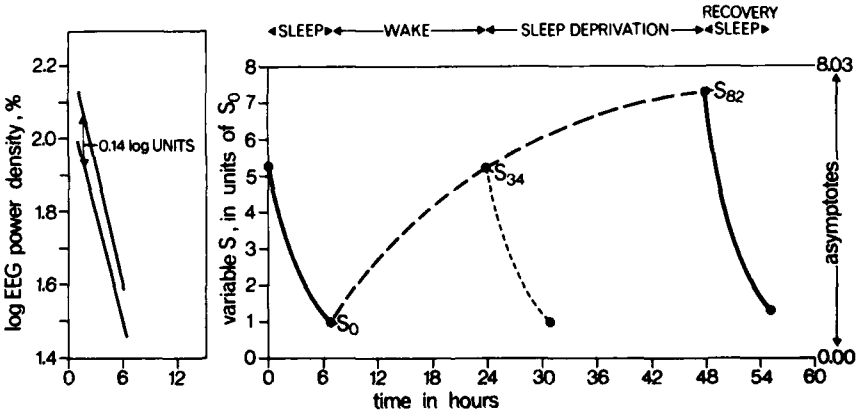


FIG. 3. Left: Data of Borbély et al. (10) on EEG power density during sleep in humans. *Lower line:* baseline night. *Upper line,* recovery night after sleep deprivation. **Right:** Derivation of the exponential rise of S during wakefulness, on the basis of the data in the left panel.

the logarithmic plot of Fig. 3, left panel. During sleep, S asymptotically approaches zero.

In a similar way, the results of sleep deprivation suggest that the buildup of S during activity can be approximated by an exponential process of the form

$$S_i = A - (r)^i \cdot (A - S_0) \tag{2}$$

where S_i is the value of S at i time units (half hours) after sleep termination. During activity, S approaches an upper asymptote A. If a normal sleep duration of 7 hr is assumed, the buildup rate r can be estimated from the breakdown rate during sleep and the increase during sleep deprivation, as follows:

S_{34} is the value of S at 34 half hours after sleep termination, that is, at sleep onset. During activity, 7 hr of breakdown are compensated for, and

$$S_{34} = [1/(0.888)^{14}] \cdot S_0 = 5.26 \cdot S_0 = A - r^{34} \cdot (A - S_0) \tag{3}$$

During 24 hr of sleep deprivation after $t = 34$, thus at $t = 82$, S is further increased by 0.14 log units [as in Borbély et al. (10), Fig. 10], such that

$$S_{82} = 1.38 \cdot 5.26 \cdot S_0 = 7.26 \cdot S_0 = A - r^{82} \cdot (A - S_0) \tag{4}$$

Subtraction of S_0 from both equations 3 and 4 yields

$$1 - r^{34} = 4.26 \cdot S_0/(A - S_0)$$

$$1 - r^{82} = 6.26 \cdot S_0/(A - S_0)$$

$$r^{34} - 1 = 0.681 \cdot (r^{82} - 1) \tag{5}$$

Numerical approximation of equation 5 gives a value of r of 0.973. This value is independent of both S_0 and A. For convenience, we have chosen a value of 1.0 for the upper asymptote A, and S_0 then equals 0.124.

Next we assume that the buildup of S is terminated by sleep onset when an upper threshold T_h is reached and that the breakdown of S during sleep is terminated when S reaches a lower threshold T_l. Thresholds T_h and T_l are, on average, symmetrically distributed around the level L. The emerging self-regulating process, in the absence of circadian modulation, is illustrated in Fig. 4. It is obvious that the periodicity of buildup (duration α) and breakdown (duration ρ) emerging depends on the threshold levels. In analogy with normal thermostats, we have a "somnostat" with a frequency depending on the settings of the upper and lower thresholds. This periodic process acts as a relaxation oscillator and is distinct from true self-sustained oscillators by having a positive rather than a negative correlation between α and ρ. For instance, incidental random elevation of T_h will lead to a longer α followed by a longer ρ.

In order to let our somnostat run in synchrony with the day–night cycle, we assume that there are circadian variations in the threshold levels T_h and T_l. We have not systematically explored the alternative possibility, i.e., that the parameters of the homeostatic process, r and d, are subject to circadian

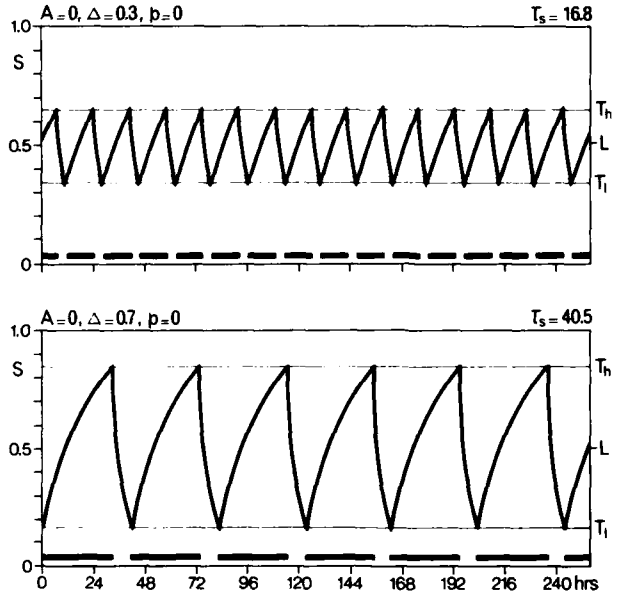


FIG. 4. Deterministic simulations of the self-regulated sleep–wake process, showing a high frequency when the distance between thresholds is small and a low frequency when the distance is large.

modulation. However, the latter seems to us the less plausible of the two alternatives, for the following reason: to cause normal daily timing of sleep, in this alternative, the buildup of S should be faster in the evening than in the morning, whereas breakdown should be slower in the evening than in the morning. This deduction is incompatible with the general proposition that the sleep EEG will reflect some aspect of the breakdown process. In our model, we have assumed that the two thresholds, T_h and T_l, are controlled simultaneously by a single circadian oscillator. It is not unthinkable that the two are controlled separately by different oscillators or by the same oscillator generating different amplitudes. However, such complications are presently uncalled for.

For further simplicity, we have assumed that the thresholds vary in sinusoidal fashion, as a generalized oscillatory pattern. A nonlinear oscillatory movement would, of course, be more realistic for biological systems, but it would not affect the qualitative predictions obtained. We do not specify how this oscillation in the sleep threshold is generated, but suggest that a circadian pacemaker such as the suprachiasmatic nucleus (45) is involved that may express itself simultaneously in other physiological variables, such as body temperature. The introduction of parallel sine waves in the thresholds forces us to specify three other parameters: period, amplitude, and phase. Of these, phase specification is

irrelevant, because we shall be discussing the steady-state behavior of the system rather than initial conditions.

For the period (τ) of the oscillation in free run, we have generally chosen 25 hr, this value being the grand average of human free-running circadian periods (53). We shall initially be concerned with free-running behavior, and later we shall simulate entrained conditions simply by setting τ at 24 hr.

We did not include in the model any feedback effect from sleep on τ. Such an effect is an essential feature in both Wever's model (53) and Eastman's model (*this volume*). The empirical evidence of the effect is not very strong. Sleep-deprivation experiments in rats, lasting either 24 hr (9) or shorter times (D. Beersma, *unpublished data*), revealed no effects on circadian period. Also, rat brain neurotransmitter receptor rhythms were unaffected by sleep deprivation (56). Circadian redistribution of activity and rest in response to restricted feeding schedules likewise leaves the circadian pacemaker essentially uninfluenced (13,25,30). On the other hand, there is some evidence for an influence of the sleep–wake cycle on the circadian oscillator in human subjects studied in isolation. When spontaneous internal desynchronization occurs, with repeated delay phase shifts of the sleep–wake cycle, the body-temperature rhythm often exhibits a shorter period (17). The average difference in τ before and during internal desynchronization is small (0.70 ± 0.38 hr) but significantly different from zero ($p < 0.001$) (53, p. 52). It is likely that in humans, studied in light, sleep behavior affects the light cycle perceived, even if it is not actively selected. Changing phase relationships between the light and the circadian oscillator in internal desynchronization may possibly affect the oscillator's period. In the DD studies (i.e., constant darkness) in rats, no such influences can be expected. Incorporating a feedback effect due to light would probably lead to a minor refinement of our model, but this is presently not essential.

Another element to be introduced in our system is stochastic variation. Again we have chosen to introduce such variation in the thresholds rather than in the variable S itself. This has intuitive appeal. It is common knowledge that the response to sleepiness may be consciously or subconsciously suppressed, e.g., by people enjoying lively social interaction late at night. On the other hand, sleep may be stimulated by circumstances such as sitting in a warm dark lecture hall while listening to a dull scientific presentation. It would be contrary to our general proposition to assume that the lively activity would suppress the sleep-regulating variable, S, rather than enhance it, or that reduced activity in the latter case would increase S. Therefore, in introducing noise into the system, we have assumed that it affects the behavioral response to fatigue rather than fatigue itself.

A final comment on the design of our model concerns the phenomenology of sleep. We have implicitly disregarded the various sleep stages. The data of Borbély et al. (10) we used concern the EEG power density during consecutive sleep cycles. The temporal distribution of this density during sleep is

homogeneous. Although there is an increasing body of knowledge regarding physiological changes [e.g., in thermoregulation (29)] associated with sleep stages, there is no evidence suggesting to us whether or not breakdown of S is restricted to specific stages. For simplicity, we have assumed that breakdown is continuous regardless of sleep stage. The parameter describing the rate of breakdown (d) was in any case derived from complete sleep periods, independent of when precisely this breakdown occurred.

DETERMINISTIC SIMULATIONS: EXPLORATION OF THE PARAMETER SPACE

The model described here was formulated as a FORTRAN algorithm, and simulations were obtained using a PDP-11 computer. The behavior of the model is fully determined by seven variable input parameters:

r = rate of buildup of S during wakefulness
d = rate of breakdown of S during sleep
Δ = distance between upper threshold (T_h) and lower threshold (T_l)
L = mean level of thresholds
τ = period of oscillation in the thresholds
A = amplitude of oscillation in the thresholds
p = standard deviation of a Gaussian $N(0,p)$ distribution from which a random variable x is drawn once per time unit and is added to both thresholds. In addition, half the deviation from the sine function in the preceding time unit is added to allow for some continuity in the threshold noise.

Of these parameters, r and d were held constant at 0.888 and 0.973, respectively, on the basis of Borbély's data (Fig. 3); τ was likewise fixed, at 25 hr. Parameter L was initially set at 0.5, such that the thresholds were symmetrically distributed with respect to the asymptotes of the S process (see Figs. 3–7). L primarily affects the α/ρ ratio of the sleep–wake periodicity, and thus, for the stochastic simulations (see Figs. 8–16), a value of $L = 0.42$ was selected to obtain more realistic sleep times of about 8 hr (see Fig. 16). We are therefore left with three free parameters: Δ, A, and p. We shall first consider the case in which $p = 0$ and discuss the effects of variations in Δ and A in the deterministic version of the model.

With $A = 0$, i.e., in the absence of circadian threshold oscillations, an increase in Δ causes a regular lengthening of the sleep–wake periodicity (Fig. 4). For other values of A, the lengthening of the sleep–wake periodicity at increasing values of Δ is, of course, observed only when the threshold oscillations do not synchronize the sleep–wake rhythm. This tendency toward longer periods can be observed even in conditions of synchronization, because the sleep episodes occur in a later phase of the threshold oscillations. For a

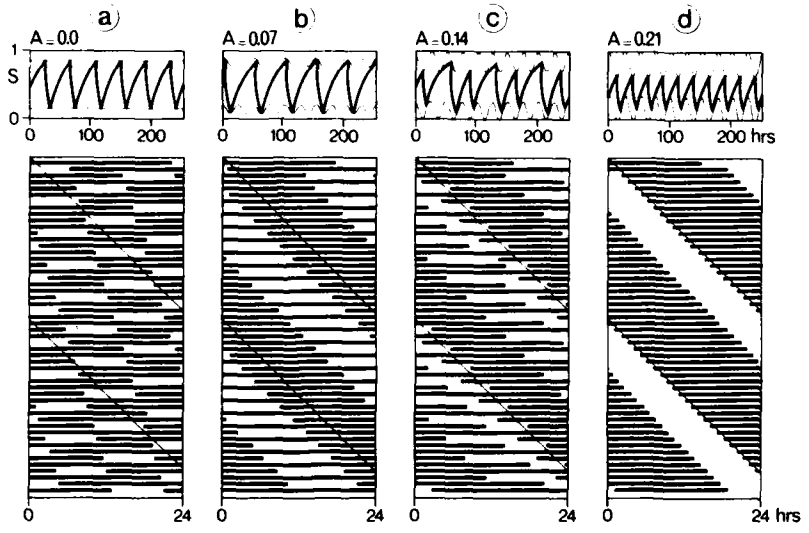


FIG. 5. Deterministic simulations showing the effects of different amplitudes of threshold oscillations (*A*) on the ensuing sleep–wake rhythm. Upper graphs show the temporal course of *S* (*heavy line*) and T_h, T_l (*thin lines*). Lower diagrams plot the same data in actogram format. Thin lines pass through consecutive zero phases (*inflection points of the upward slope*) of the threshold oscillation.

fixed value of Δ (set at 0.7 in the example of Fig. 5) and increasing values of A, there is an increasing tendency for the sleep–wake periodicity to lock onto the circadian oscillation, either by frequency demultiplication (Fig. 5b,c) or in 1:1 synchrony (Fig. 5d). Figure 5b shows a case described as *circabidian* periodicity in the literature (31,53). The average period τ_s of the sleep–wake cycle was 50 hr in Fig. 5b, as compared with 25 hr in Fig. 5d and 36.7 hr (2:3 synchronization) in Fig. 5c.

A more complete exploration of the effects of variations in Δ and A on the average period is summarized in Fig. 6. When A is close to zero, a wide range of periods (τ_s = 5.0–82.9 hr) is obtained by varying Δ between 0.05 and 0.95. When A is large, only a few periods of the sleep–wake rhythm can be observed, centered around 12.5, 16.7, 25, 36.7, and 50 hr.

Whereas circadian rhythms typically continue and reveal their own natural periods when released from entrainment into constant conditions, their amplitudes (when measured) usually are lower in free run than in entrainment. This effect is well documented in numerous plant circadian rhythms (14), and it also seems widespread in rhythms of animal behavior and physiology, e.g., in chicken brain temperatures (5), chaffinch oxygen consumption (43), and body temperatures in rats (21,46) and squirrel monkeys (24). In humans, also, there is evidence that the amplitudes of various physiological oscillations decrease in

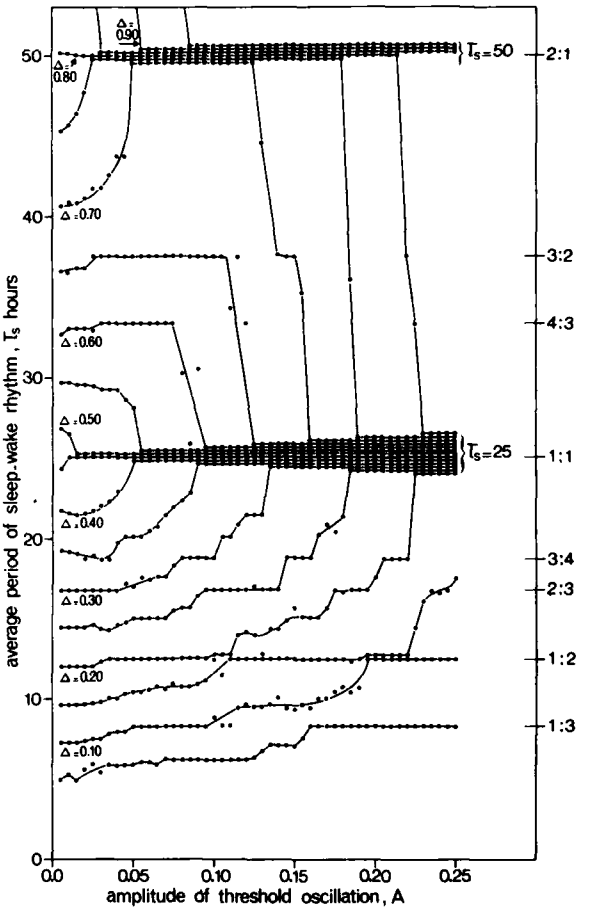


FIG. 6. Average period τ_s of the sleep–wake rhythms as a function of A, for 19 values of Δ. Ratios to the right show some preferred patterns of synchrony and frequency multiplication of the system. When A is large, only a few preferred frequencies occur; when A is small, τ_s can have almost any value.

association with the change from entrained to free-running conditions, e.g., in body temperature and plasma cortisol (49,53). There is thus ample basis for proposing that the circadian oscillation in the S thresholds may decrease after entry to free run. There are various instances in Fig. 6 where, with constant Δ, a reduction in A leads to rather sudden changes in the resulting period, τ_s. However, each point in Fig. 6 is obtained from a simulation with A and Δ kept constant over 50 threshold periods. The results of this static approach do not necessarily predict what happens with a changing value of A. In simulations

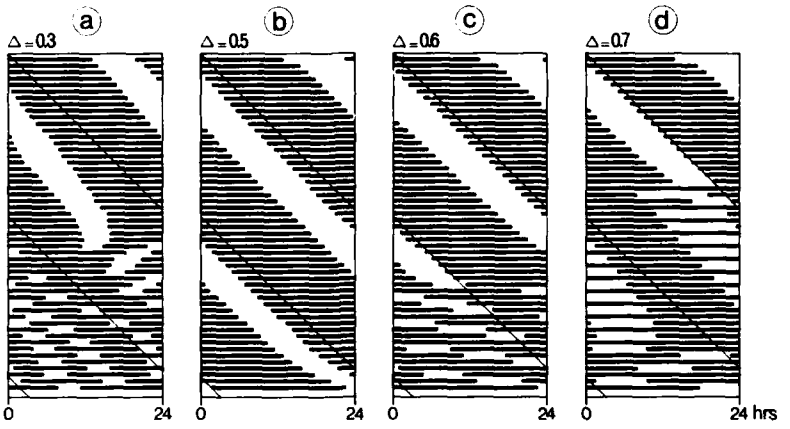


FIG. 7. Deterministic simulation with A decreasing linearly from 0.25 to 0 in the course of 50 cycles for various values of Δ. Note the abrupt occurrence of internal desynchronization in examples a, c, and d.

where A is gradually reduced by 0.005 S units per cycle, as in Fig. 7, we observe an initial periodicity of 25 hr, which changes gradually until a sudden transition to 2:3 or 1:2 synchrony occurs. We therefore propose that it is the reduction in threshold amplitude in isolated conditions that has led in some individuals to "internal desynchronization" of sleep and wakefulness from other physiological rhythms. This explains why such desynchronization occurs more frequently in free run than in normal life. A gradual reduction would also explain why desynchronization is typically preceded by a number of days of internally synchronized free run.

STOCHASTIC SIMULATIONS: PRECISION AND INTERNAL SYNCHRONY

The model in the deterministic formulation does not account for the irregularity of the patterns of internal desynchronization observed in the studies of Wever (53) and Czeisler (16). The regularity of the circabidian case, with phase jumps of the sleep–wake rhythm occurring every other circadian cycle, is exceptional. More commonly, such phase jumps are observed in variable intervals of 2 to 10 cycles [e.g., Wever (53), Figs. 30, 32, and 36]. Such unpredictability is an essential aspect of the behavior of the system, and it obviously calls for some sort of stochastic variation. Furthermore, some interesting analyses of variability in human sleep–wake rhythms have been reported (2). It is appropriate to ask if observed patterns of variability may be a consequence of circadian gating of a homeostatic process as we propose.

For purposes of simulation, we introduced a fixed noise parameter ($p = 0.05$) in the system. Together with the coarse time scale we used (using time units of half hours), this led to reasonable variance in the cycle length (standard deviations of τ_s usually between 1 and 3 hr in conditions of synchrony). The same variation affected upper and lower thresholds. This is probably somewhat unrealistic. It is likely that conditions affecting the thresholds are much more constant during sleep than during activity. It was, however, not quite clear to us which noise values had to be selected for the upper and lower thresholds, and in this chapter we restrict ourselves to equal noise. We have further chosen $A = 0.12$, $L = 0.42$, and $\Delta = 0.5$ as standard parameter values in entrained conditions, in order to obtain an optimal fit to the available human data (Fig. 16).

The introduction of noise into the threshold not only leads to variations in onset and end of activity during synchronized free run but also leads to a certain amount of unpredictability in the occurrence of phase jumps, as a sample simulation in Fig. 8 shows. Such phase jumps can occur with any average frequency, depending on the values of A and Δ. Some preferred frequencies are 0 ("internal synchrony") and once per two cycles ("circabidian pattern"). In an infinitely long simulation, any average frequency of the ensuing sleep–wake rhythm may be realized.

More informative of the nature of the underlying process are the distributions of sleep duration and the phases of sleep onset relative to the circadian system. Such data from experiments with human subjects in isolation have been thoroughly analyzed by Zulley (57), as summarized in Fig. 9. When subjects showing internal desynchronization are considered, there is a clear bimodality in the time of spontaneous sleep onset relative to the minimum of the circadian cycle of body temperature. A peak in sleep-onset times is observed 2 to 3 hr

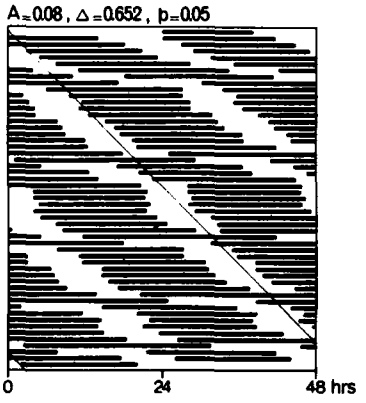


FIG. 8. Stochastic simulation, plotted in double actogram format. Note the irregular occurrence of phase jumps.

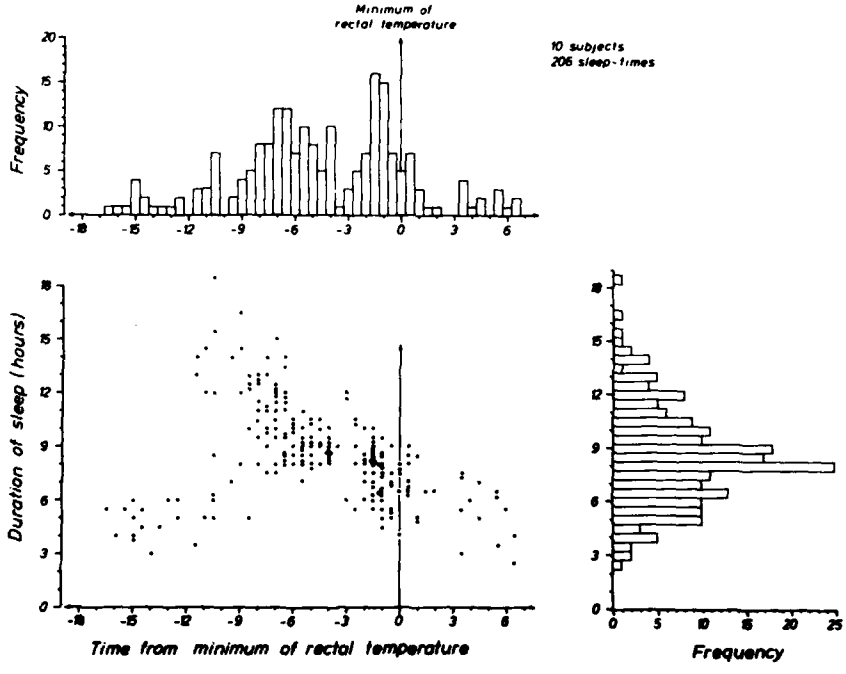


FIG. 9. Distributions of sleep onsets and sleep durations in 10 human subjects showing interval desynchronization. Data from J. Zulley. (From Czeisler et al., ref. 17a, with permission.)

before the rectal-temperature minimum. The subsequent duration of sleep is typically in the range of 6 to 9 hr and corresponds to the normal internally synchronized situation (58). A second concentration of sleep onsets occurs at approximately 7 hr before the minimum of rectal temperature and is followed by considerably longer sleep (8–12 hr). Infrequent sleep onsets occurring between 9 and 12 hr before the rectal-temperature minimum are followed by either short (ca. 6 hr) or very long (12–16 hr) sleep lengths.

This pattern, confirmed by Czeisler et al. (17), is also observed in a series of simulations such as the one presented in Fig. 8. In all of these, A was set at 0.08, i.e., slightly below the value of 0.12 chosen for normal entrained conditions, to allow for a one-third decrease in amplitude due to free run; Δ was varied above the standard value of 0.5 to represent a sample of parameter values with increased tendency toward internal desynchronization. The resulting distributions of sleep-onset times with respect to the threshold oscillation minimum and the corresponding sleep durations (Fig. 10) show reasonable agreement with the experimental data. There are, however,

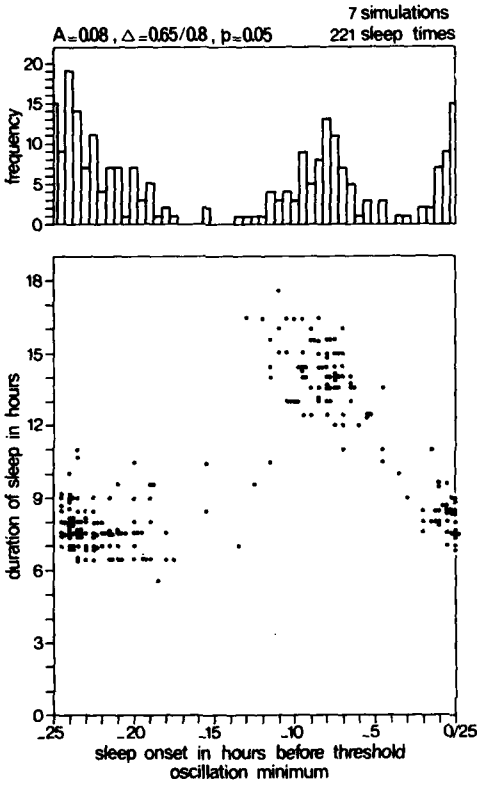


FIG. 10. Distribution of sleep onsets and sleep durations in simulations with various Δ values leading to internal desynchronization.

quantitative differences. The distributions are slightly shifted to the right, such that the main sleep-onset peak occurs shortly after, instead of before, the oscillation minimum. However, there is no reason to assume that the temperature curve will precisely match the threshold curve, especially because body temperature itself is directly influenced by sleep and activity [e.g., Aschoff and Wever (6), Fig. 4]. The early sleep-onset peak is narrower than in the experimental data (Fig. 9), and thereby the gap between the peaks is wider. This may be related to ambiguities in determining the temperature minimum and, in addition, may be the result of our use of a sine wave instead of a more realistic waveform of the thresholds, i.e., one with a flatter drop phase, where sleep onset occurs. Sleep lengths are somewhat long for the early sleep onsets. This suggests that a distorted waveform with a steeper rise of the threshold, which would shorten the sleep, would produce a better fit to the data. Later we shall indicate ways of experimentally approaching the precise shape of the threshold waveform.

Also in the state of internal synchrony the introduction of noise in the system

has some interesting consequences for the patterns of variability and precision of the system. We have not yet fully explored all of these, but we wish to outline how the system may give rise to the negative serial correlations (r_s) that have been considered indicative of the oscillatory origin of the sleep–wake rhythm (53, p. 71).

In human subjects, the lengths of consecutive sleep–wake cycles are in general negatively correlated with each other (53), as also in other circadian rhythms (41). A series of simulations with standard parameter values, but with τ varying between 23 and 27 hr, gave about the same average values of r_s, when based on activity onset (−0.31) and activity end (−0.41). These values compare well with Wever's empirical value of −0.40, valid for both (Fig. 11).

The match of the correlations between α and ρ with empirical data is much less good. Only in a qualitative sense does our model, at least with the standard parameter values, match the empirical facts that α is negatively correlated with the following ρ and that α is not significantly correlated with the preceding ρ. But in the human data, the correlation is much stronger (−0.52 as compared with −0.25). Again, this may indicate where our model is incorrect and in need of improvement.

The correlations between α and ρ depend on (a) the slope of the threshold where it is intersected by the S process and (b) the amount of noise in the threshold. The first effect may be illustrated by the way in which the two correlation coefficients depend on the average phase of sleep onset (Fig. 12). These simulations used slightly different parameter values, and variations in sleep-onset phase were created by letting Δ vary over a large range. They therefore do not match any human data, but merely serve to illustrate the argument that with a later phase of sleep onset, α is more negatively correlated with the following $\rho[r(\alpha,\rho)]$ and less negatively or more positively with the preceding $\rho[r(\rho,\alpha)]$. Thus, the generally late phase of sleep onset before the threshold oscillation minimum leads to a more negative value of $r(\alpha,\rho)$. The steeper the slope at sleep onset, the more precise is sleep-onset time, and the

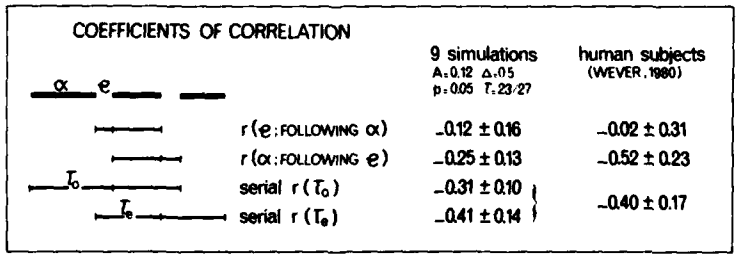


FIG. 11. Comparison of coefficients of correlation between α and ρ and between consecutive τ_s measurements, obtained in human experiments and by simulation.

less negative $r(\alpha,\rho)$ will become. In contrast, the steeper the slope at the end of sleep, the more precise is the time of sleep end, and the more negative $r(\alpha,\rho)$ becomes. A distortion of the threshold waveform with a steeper rise and less steep drop would cause $r(\alpha,\rho)$ to be more negative and bring $r(\rho,\alpha)$ closer to zero, thus leading to a better match with Wever's empirical data.

The second factor involved is the amount of noise in the two thresholds. In our model we used the same noise level in both thresholds, being aware that this is unrealistic. The standard deviations of τ_s when based on activity onset or activity end are generally equal (Fig. 13) for standard parameters (with A varying between 0.12 and 0.06 to allow for some reduction of amplitude in free run). For comparison, data presented by Aschoff et al. (2) are included in Fig. 13, showing variation in activity end (sleep onset) increased about 1.5 times relative to variation in activity onset. This suggests that a better fit would be obtained by introducing more noise in the upper threshold than in the lower threshold in our model. Such differential variation would at the same time lead to more negative $r(\rho,\alpha)$. It is noteworthy that the relative amounts of variation

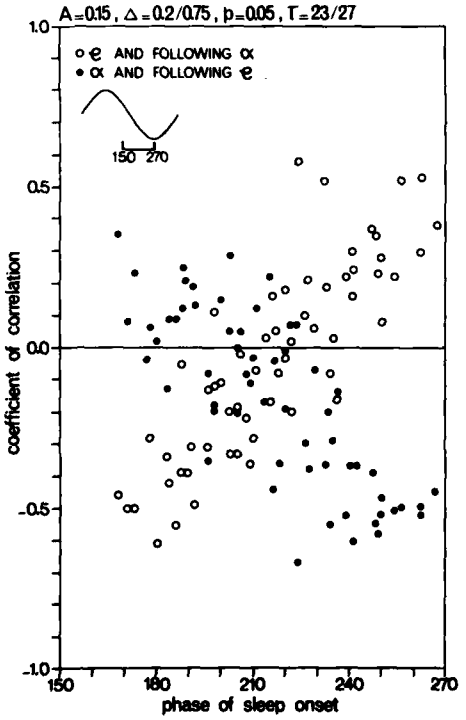


FIG. 12. Coefficients of correlation between α and ρ as a function of the average phase of sleep onset in a series of simulations in which Δ and τ are varied.

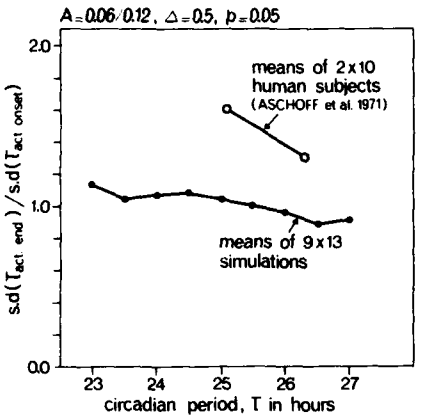


FIG. 13. Ratio of standard deviations of τ_s based on activity end and activity onset, obtained during internal synchrony in human experiments and by simulation.

in cycle lengths based on onset and end of activity change with τ in the same manner as observed in human data (Fig. 13). In our simulations this is due to changes in the phase of the sleep–wake process with changes in τ. Aschoff et al. (2) attributed the effect to a change in form factor of the oscillation in their model.

Clearly, the analysis of variation in the system may suggest some improvements in the model, but we postpone a more exhaustive analysis until more empirical data have been incorporated in it.

MANIPULATION OF THE SLEEP–WAKE SYSTEM

So far we have discussed only the spontaneous behavior of the sleep–wake system in conditions of isolation from exogenous time cues. There was a reasonable correspondence between the behavior of our model and the main phenomena of human sleep and wakefulness. However, the model should be put to the test by means of experimental manipulation of the system, rather than by merely phenomenological similarities between simulation and observation.

Essentially, such experiments would involve manipulations of the thresholds rather than of the *S* renewal process itself. We presume that the upper threshold can be strongly affected by prevailing conditions and by conscious decisions. It may be raised during forced nocturnal shiftwork, during sleep-deprivation experiments, during lively social interaction at a late-night party. Warmth, darkness, silence, a comfortable bed or chair, and absence of intellectual or physical activity all represent conditions characterizing low upper and lower thresholds. The lower threshold may be raised and early awakening induced by noise (alarm clock), cold, bright light, etc.

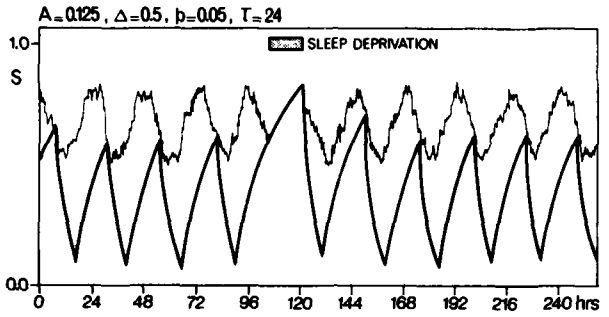


FIG. 14. Example of the effect of sleep deprivation for 16.5 hr on the temporal course of *S*. The lower threshold, running parallel at a distance of 0.5 below the upper threshold, is omitted for clarity.

Only sleep deprivation has so far been systematically explored. We shall compare its effect with temporally raising the upper threshold in our model. An example of a simulation, with 16.5 hr of sleep deprivation (SD) on night 5, is shown in Fig. 14. The *S* value continues to rise until the deprivation is stopped, and a slightly lengthened recovery sleep follows. However, the sleep length following sleep deprivation is not always lengthened in such simulations. The lower panel of Fig. 15 plots recovery sleep as a function of prior SD. Initially, recovery sleep gradually shortens with increasing SD, jumping to a maximum after 16 hr of SD, and then gradually returning to only slightly elevated recovery sleep times after 24 hr of SD.

Careful data on the duration of voluntary sleep following different lengths of SD have recently been collected by Åkerstedt and Gillberg (1,26). The original data were kindly sent to us by Dr. Åkerstedt, and we plotted the individual values in the upper panel of Fig. 15. The same general shape of the sleep lengths as a function of prior SD is observed. At least in three of the six subjects a steep rise was seen, either between 12 and 16 or between 16 and 20 hr of SD.

We would not anticipate a better match with our simulations, because the assumption of a sine wave as the threshold waveform is unrealistic. In fact, the data of Åkerstedt and Gillberg (1) may be used to obtain a crude estimate of the waveform of the lower threshold. This is attempted in Fig. 16. Normal bedtime for the Stockholm subjects was hour 23:00 and normal sleep length 8 hr. This fully defines the values of *S* at any time of day in our model. Assuming that *S* continues to rise exponentially during sleep deprivation, the values labeled 3, 7, 11, etc., in Fig. 16 are obtained at the respective clock times. Because average voluntary sleep lengths following SD were measured, the values of *S* after an exponential drop during recovery sleep can be derived. A line fitted by eye through these points yields the average lower-threshold curve during bed rest. This was extended left and right with a periodicity of 24 hr and reproduced

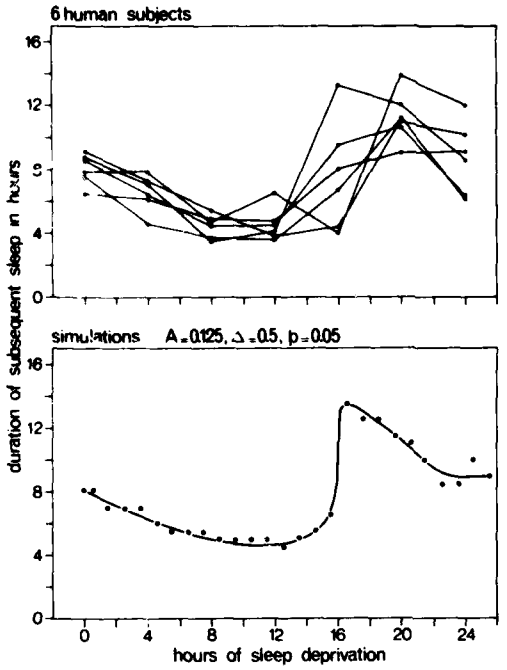


FIG. 15. Effect of different durations of sleep deprivation on the subsequent duration of voluntary sleep. **Upper panel:** Experimental data (courtesy of Dr. Åkerstedt). **Lower panel:** Simulation results.

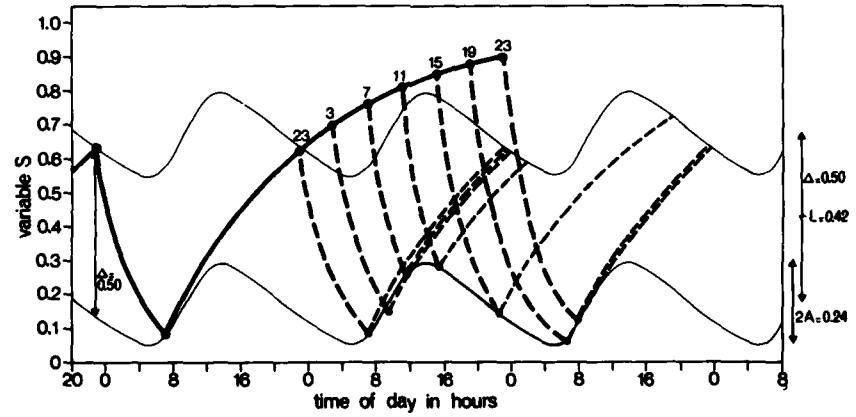


FIG. 16. Reconstruction of the lower threshold using the data on sleep deprivation of Åkerstedt and Gillberg (1). Derived parameter values indicated to the right were used in most stochastic simulations, albeit in association with sine-wave threshold curves.

upward to pass through the point at hour 23:00 to obtain a first estimate of the upper-threshold curve under the assumption that the two run parallel.

The lower-threshold curve, as thus constructed from the Stockholm data, is markedly skewed, with a steep rise followed by a less steep fall. On several occasions before, we had other reasons to assume a steeper rise [on the basis of Zulley's data (57) on spontaneous sleep lengths in isolation, Wever's (53) measurements of $r(\alpha,\rho)$ and $r(\rho,\alpha)$, and calculations of cycle-to-cycle variance by Aschoff et al. (2)]. There is thus good reason to reject the sine-wave model. In a subsequent publication (D. Beersma and S. Daan, *in preparation*) we shall analyze the model using the threshold curve empirically derived as in Fig. 16.

Processes other than SD by which the system can be manipulated include phase shifting of the circadian oscillator, such as in transmeridian flights. The transients observed in such experiments (3) seem to last longer than can be accounted for by the gradual regaining of synchrony of the sleep–wake rhythm in response to an instantaneously reset oscillator. As long as no straightforward assumptions can be made regarding the kinetics of oscillator resetting, we do not find it useful to study this kind of manipulation in model simulations at this time.

PREDICTIONS AND PERSPECTIVE

Because the model discussed generates precise predictions of the temporal course of sleep and wakefulness, it is possible to test various elements in it in quantitative detail.

The first type of prediction concerns Borbély's basic hypothesis that the EEG power density reflects the level of the sleep–regulating variable. If this is correct, it should be possible to estimate precisely the kinetics of the S buildup process. This can be done by recording initial EEG power densities of subjects who are allowed to fall asleep in strictly controlled conditions at different times of day and after different durations of sleep deprivation. Also, it should be possible to predict the initial EEG power density after any exotic forced-sleep schedule (e.g., 7:00 wake up; 23:00 sleep onset; 1:00 wake up; 12:00 sleep onset). One example of such predictions is implied in Fig. 16. Subjects in Åkerstedt and Gillberg's (1) study when sleep-deprived for 16 hr (until 15:00), slept until 22:30. These subjects should also show a higher EEG power density than normal on the second recovery night when starting sleep the next day at 19:30. In contrast, after 24 hr of sleep deprivation the EEG is back to normal on the second recovery night (10).

The second type of prediction concerns the lower or wake-up threshold. It is crucial to the predictive value of the model that this is independent of the actual value of S. It was shown by Åkerstedt and Gillberg (1, Fig. 2) that SD lasting until 3, 7, 11, and 15:00 led to high terminal SWS percentages, whereas low

terminal SWS percentages were seen after normal sleep and SD lasting until 19 and 23:00. Because EEG power density is determined mainly by slow-wave sleep, this suggests that terminal S values do indeed vary in circadian manner as suggested by Fig. 16. It should be possible to collect more empirical points for the lower-threshold curve by applying other, more complex sleep–wake schedules.

A much more difficult task is to determine the upper threshold. This curve is presumably subject to much larger variations due to external conditions and to conscious decisions of subjects. Only carefully controlled conditions, such as with continuous bed rest, can allow quantification of at least the falling part of the curve. Figure 16 suggests that studying spontaneous sleep behavior in the second night following SD experiments may be useful in this context. The curve may, to some extent, be related to subjective sleepiness ratings, because sleepiness at bedtime in the study of Åkerstedt and Gillberg (1, Fig. 3) was lowest at bedtimes 23, 11, and 15, where also in our computations (Fig. 16) S should be considerably less above threshold than at 3, 7, 19, and 23.

Flexibility of the thresholds in response to conditions may also be involved in adjustments of sleep–wake periodicity to conscious or subconscious habits, to shiftwork, and to Wever's forced-sleep schedules (53, Figs. 86–94) by applying strong light–dark Zeitgebers. The consequences of such exogenous schedules for the sleep–wake process can of course be easily simulated with the model presented, but we postpone this until a more realistic formulation (not using sine waves) is available. Such a formulation may have to include bimodality of the threshold curve to account for a midday drop in sleep latency and afternoon napping behavior (15).

Finally, our model may find application in understanding anomalies of the circadian system such as occur in some depressive conditions. The usual reduction in slow-wave sleep in such patients (27) is suggestive of a number of possible anomalies in the system: An advanced phase of threshold oscillation [corresponding to a hypothesis of Wehr et al. (50)], lowered thresholds, or a reduced rate of S buildup would all lead to shorter and shallower sleep. Precise consequences of such deviations remain to be explored.

Finally, we wish to comment on some general implications of gating systems as proposed here for human sleep and wakefulness.

The model presented is certainly not correct in great quantitative detail. It shows, however, that the main characteristics of human sleep–wake rhythms can be produced without the help of a second oscillator or pacemaker. Even the negative correlation between α and ρ, seen by Wever (53, p. 71) as evidence of the oscillatory origin of the sleep–wake rhythm, is likewise produced by our relaxation system. We have expounded the view that sleep may have evolved as a means of allowing organisms to exploit their nervous tissues during part of the day–night cycle more intensively than would be possible on a permanent basis. Sleep would then provide the necessary recovery (whatever its nature) during the other, adverse part of the cycle. Such a process might, of course, be

controlled by a separate central pacemaker, as suggested by Aschoff and Wever (6), or directly by the central pacemaker (22). The presence of a single circadian pacemaker that normally gates the sleep–wake process such that sleep occurs at the adaptively correct time of day is the more parsimonious hypothesis, and it also has intuitive appeal. Such a system would allow for the flexibility needed by animals in nature. Sometimes food may be more abundant in that part of the day–night cycle in which sleep normally occurs. A host of experiments show that activity rhythms are rapidly adjusted to such unusual circumstances (18), while the central timer probably keeps running under LD control. Such force of external conditions and daily habits may affect the threshold system such that sleep occasionally shifts to unusual parts of the cycle (the night in rats fed only in daytime; the day in workers on a night shift). For optimal and flexible timing, a precise central oscillator is required, exerting gentle control over the processes it gates.

ACKNOWLEDGMENTS

The idea of our approach originated at the Ringberg Conference on Structure and Physiology of Vertebrate Circadian Systems in response to stimulating papers by Drs. Eastman, Borbély, and Zulley. Later, suggestions arose from discussions with Drs. Aschoff, Groos, v. d. Hoofdakker, and Wirz-Justice. Drs. Åkerstedt, Aschoff, and Borbély were so kind as to put their experimental data at our disposal. We are further grateful to Dick Visser for drawing the figures and to Mrs. H. Lochorn-Hulsebos for typing the manuscript.

REFERENCES

1. Åkerstedt, T., and Gillberg, M. (1981): The circadian variation of experimentally displaced sleep. *Sleep*, 4:159–169.
2. Aschoff, J., Gerecke, U., Kureck, A., Pohl, H., Rieger, P., von Saint Paul, U., and Wever, R. (1971): Interdependent parameters of circadian activity rhythms in birds and man. In: *Biochronometry*, edited by M. Menaker, pp. 3–27. National Academy of Sciences, Washington, D. C.
3. Aschoff, J., Hoffmann, K., Pohl, H., and Wever, R. (1975): Re-entrainment of circadian rhythms after phase-shifts of the Zietgeber. *Chronobiologia*, 2:23–78.
4. Aschoff, J., von Goetz, C. H., and Honma, K. (1983): Restricted feeding in rats: effects of varying feeding cycles. *Z. Tierpsychol. (in press)*.
5. Aschoff, J., and von Saint Paul, U. (1973): Brain temperature as related to gross motor activity in the unanaesthetized chicken. *Physiol. Behav.*, 10:529–533.
6. Aschoff, J., and Wever, R. (1981): The circadian system of man. In: *Handbook of Behavioral Neurobiology, Vol. 4, Biological Rhythms*, edited by J. Aschoff, pp. 311–322. Plenum, New York.
7. Berger, R. J. (1975): Bioenergetic functions of sleep and activity rhythms and their possible relevance to aging. *Fed. Proc.*, 34:97–102.
8. Borbély, A. A. (1980): Sleep: Circadian rhythm versus recovery process. In: *Functional States of the Brain: Their Determinants*, edited by M. Koukkow, D. Lermann, and J. Angst, pp. 151–161. Elsevier, Amsterdam.
9. Borbély, A. A. (1982): Sleep regulation: Circadian rhythm and homeostasis. In: *Sleep: Clinical and Experimental Aspects*, edited by D. Ganten and D. Pfaff, pp. 83–103. Springer, Berlin, Heidelberg, New York.

10. Borbély, A., Baumann, F., Brandeis, D., Strauch, I., and Lehrmann, D. (1981): Sleep deprivation: Effect on sleep stages and EEG power density in man. *Electroencephalogr. Clin. Neurophysiol.*, 51:483–493.
11. Borbély, A., and Tobler, I. (1980): The search for an endogenous 'sleep substance.' *Trends Pharmacol. Sci.*, 2:356–358.
12. Borbély, A. A., Tobler, I., and Groos, G. (1982): Sleep homeostasis and the circadian sleep-wake rhythm. In: *New Perspectives in Sleep Research*, edited by M. H. Chase, Spectrum, Jamaica, New York.
13. Boulos, Z., Rosenwasser, A. M., and Terman, M. (1980): Feeding schedules and the circadian organization of behavior in the rat. *Behav. Brain Res.*, 1:39–65.
14. Bünning, E. (1977): *Die physiologische Uhr*, 3rd ed. Springer, Berlin.
15. Carskadon, M. A. (1983): Daytime sleep tendency. Implications for napping. In: *Sleep Research*, Vol. 10, edited by M. H. Chase, D. F. Kripke, and P. L. Walter. Brain Information Service/Brain Research Institute, Los Angeles (*in press*).
16. Czeisler, C. (1978): Human circadian physiology: Internal organization of temperature, sleep-wake and neuroendocrine rhythms monitored in an environment free of time cues. Ph.D. dissertation, Stanford University.
17. Czeisler, C., Weitzman, E. D., Moore-Ede, M. C., Zimmerman, J. C., and Kronauer, R. S. (1980): Human sleep: Its duration and organization depend on its circadian phase. *Science*, 210:1264–1267.
17a. Czeisler, C. A., Zimmerman, J. C., Ronda, J. M., Moore-Ede, M. C., and Weitzman, E. D. (1980): Timing of REM sleep is coupled to the circadian rhythm of body temperature in man. *Sleep*, 2:329–346.
18. Daan, S. (1981): Adaptive daily strategies in behavior. In: *Handbook of Behavioral Neurobiology, Vol. 4: Biological Rhythms*, edited by J. Aschoff, pp. 275–298. Plenum, New York.
19. Daan, S., and Berde, C. (1978): Two coupled oscillators: Simulations of the circadian pacemaker in mammalian activity rhythms. *J. Theor. Biol.*, 70:297–313.
20. Davis, F. C., and Menaker, M. (1980): Hamsters through time's window: Temporal structure of hamster locomotor activity. *Am. J. Physiol.*, 239:R149–R155.
21. Eastman, C. I. (1980): Circadian rhythms of temperature, waking, and activity in the rat. Ph.D. thesis, University of Chicago.
22. Enright, J. T. (1980): *The Timing of Sleep and Wakefulness*. Springer-Verlag, Berlin.
23. Fencl, V., Koski, G., and Pappenheimer, J. R. (1971): Factors in cerebrospinal fluid from goats that affect sleep and activity in rats. *J. Physiol. (Lond.)*, 216:565–589.
24. Fuller, C. A., Sulzman, F. M., and Moore-Ede, M. C. (1978): Thermoregulation is impaired in an environment without circadian time cues. *Science*, 199:794–796.
25. Gibbs, F. B. (1979): Fixed interval feeding does not entrain the circadian pacemaker in blind rats. *Am. J. Physiol.*, 236:R249–R253.
26. Gillberg, M. and Åkerstedt, T. (1983): Body temperature and sleep at different times of day. *Sleep* (*in press*).
27. Gillin, J. C., Duncan, W. C., Murphy, D. L., Post, R. M., Wehr, T. A., Goodwin, F. K., Whyatt, R. J., and Bunny, W. E. (1981): Age-related changes in sleep in depressed and normal subjects. *Psychiatry Res.*, 4:73–78.
28. Hartmann, E. L. (1973): *The Functions of Sleep*. Yale University Press, New Haven.
29. Heller, H. C., and Glotzbach, S. F. (1977): Thermoregulation during sleep and hibernation. *Int. Rec. Physiol. Environmental Physiol.*, 15:147–187.
30. Honma, K., von Goetz, C. H., and Aschoff, J. (1983): Effects of restricted daily feeding on free running circadian rhythms in rats. *Physiol. Behav.* (*in press*).
31. Jouvet, M., Mouret, J., Chouvet, G., and Siffre, M. (1974): Toward a 48-hour day: Experimental bicircadian rhythm in man. In: *The Neurosciences: Third Study Program*, edited by F. A. Schmitt and G. Worden, pp. 491–497. M.I.T. Press, Cambridge, Mass.
32. Kronauer, R. E., Czeisler, C. A., Pilato, S. F., Moore-Ede, M. C., and Weitzman, E. D. (1982): Mathematical model of the human circadian system with two interacting oscillators. *Am. J. Physiol.*, 242:R3–R17.
33. Meddis, R. (1975): On the function of sleep. *Anim. Behav.*, 23:676–691.
34. Meddis, R. (1977): *The Sleep Instinct*. Routledge & Kegan Paul, London.

35. Mistlberger, R., Waldenar, W., and Rechtschaffen, A. (1981): Effect of suprachiasmatic nuclei lesions on rebounds from sleep deprivation. In: *Sleep Research*, Vol. 10, edited by M. H. Chase, D. F. Kripke, and P. L. Walter. Brain Information Service/Brain Research Institute, Los Angeles (*in press*).
36. Nagasaki, H., Kitahama, K., Valatx, J. L., and Jouvet, M. (1980): Sleep-promoting effect of the sleep-promoting substance (SPS) and delta-sleep inducing peptide (DSIP) in the mouse. *Brain Res.*, 1972:276–280.
37. Pappenheimer, J. R., Koski, G., Fencl, V., Karnovsky, M. L., and Krueger, J. (1975): Extraction of sleep-promoting factor S from cerebrospinal fluid and from brains of sleep-deprived animals. *J. Neurophysiol.*, 38:1299–1311.
38. Pavlidis, T. (1967): A model for circadian clocks. *Bull. Math. Biophys.*, 29:781–791.
39. Pavlidis, T. (1973): *Biological Oscillators: Their Mathematical Analysis.* Academic Press, New York.
40. Pittendrigh, C. S. (1960): Circadian rhythms and the circadian organization of living systems. *Cold Spring Harbor Symp. Quant. Biol.*, 25:159–182.
41. Pittendrigh, C. S., and Daan, S. (1976): A functional analysis of circadian pacemakers in nocturnal rodents. I. Stability and lability of spontaneous frequency. *J. Comp. Physiol.*, 106:233–252.
42. Pittendrigh, C. S., and Skopik, S. D. (1970): Circadian systems. V. The driving oscillation and the temporal sequence of development. *Proc. Natl. Acad. Sci. USA*, 65:500–507.
43. Pohl, H. (1970): Zur Wirkung des Lichtes auf die circadiane Periodik des Stoffwechsels und der Aktivität beim Buchfinken, *Z. Vergl. Physiol.*, 66:141–163.
44. Rechtschaffen, A. D. (1979): The function of sleep: Methodological issues. In: *The Functions of Sleep*, edited by R. Drucker-Colin, M. Shkurovich, and M. B. Sterman, pp. 1–18. Academic Press, New York.
45. Rusak, B., and Zucker, I. (1979): Neural regulation of circadian rhythms. *Physiol. Rev.*, 59:449–526.
46. Spencer, F., Shirer, H. W., and Yochim, J. M. (1976): Core temperature in the female rat: Effect of pinealectomy or altered lighting. *Am. J. Physiol.*, 231:355–360.
47. Webb, W. B. (1971): Sleep behavior as a biorhythm. In: *Biological Rhythms and Human Performance*, edited by P. Colquhoun, pp. 149–177. Academic Press, London.
48. Webb, W. B. (1979): Theories of sleep functions and some clinical implications. In: *The Function of Sleep*, edited by R. Drucker-Colin, M. Shkurovich, and M. B. Sterman, pp. 19–35. Academic Press, New York.
49. Weitzman, E. D., Czeisler, C. A., and Moore-Ede, M. C. (1979): Sleep-wake, neuroendocrine and body temperature circadian rhythms under entrained and non-entrained (freerunning) conditions in man. In: *Biological Rhythms and Their Central Mechanism*, edited by M. Suda, O. Hayaishi and H. Nakagawa, pp. 199–227. Elsevier/North Holland, Amsterdam.
50. Wehr, T. A., Wirz-Justice, A., Goodwin, F. K., Duncan, W., and Gillin, J. C. (1979): Phase advance of the circadian sleep-wake cycle as an antidepressant. *Science*, 206:710–713.
51. Wever, R. (1965): A mathematical model for circadian rhythms. In: *Circadian Clocks*, edited by J. Aschoff, pp. 47–63. North Holland, Amsterdam.
52. Wever, R. (1967): Ueber die Beeinflüssung der circadiane Periodik des Menschen durch schwache elektromagnetischen Felder. *Z. Vergl. Physiol.*, 56:111–128.
53. Wever, R. (1979): *The Circadian System of Man.* Springer-Verlag, Berlin.
54. Winfree, A. T. (1970): An integrated view of resetting of a circadian clock. *J. Theor. Biol.*, 28:327–374.
55. Winfree, A. T. (1980): *The Geometry of Biological Time.* Springer-Verlag, Berlin.
56. Wirz-Justice, A., Tobler, I., Borbély, A., Kafka, M. S., Naber, D., Marangos, P., and Wehr, T. A. (1981): Sleep deprivation: Effects on circadian rhythms of rat brain neurotransmitter receptors. *Psychiatry Res.*, 5:67–76.
57. Zulley, J. (1979): *Der Einfluss von Zeitgebern auf den Schlaf des Menschen.* Rita G. Fischer Verlag, Frankfurt.
58. Zulley, J., Wever, R., and Aschoff, J. (1981): The dependence of onset and duration of sleep on the circadian rhythm of rectal temperature. *Pflüger Arch.*, 391:314–318.

DISCUSSION

Dr. Moore-Ede: Let me start off by clarifying one point. This is essentially a two-oscillator model. One of them really is a relaxation oscillator with a buildup of a factor which then dissipates at a threshold. And the other is a nonrelaxation oscillator. Thus, would it be fair to characterize your system as a two-oscillator rather than a one-oscillator model?

Dr. Daan: The sleep–wake process I assume would be periodic, even if you had no circadian inputs. However, I do not want to be drawn into a discussion as to whether this thing is an oscillator or not. It does not have the usual characteristics of what might be considered an oscillator. However, it is certainly periodic.

Dr. Winfree: I would like to just say that this is the first I have heard of this work, and I think it is a terrific example of modeling. I have a question. One of the things that has been troubling me is coming to terms with how sleep onset might be timed. I have detected an irregularity, therefore, in the data sets that I have not yet really come to grips with. For example, I have thought of the phase of sleep onset as a function of the phasing of the previous sleep onset or the previous wake onset or both. What we must have is a believable determination of sleep onset.

Dr. Daan: Sleep onset in my model is greatly influenced by external conditions and by conscious decisions. And so the system muffles every variation that might be prior to the timing of sleep. But occasionally you get effects where the organism is isolated from time cues, and sleep onset phase will more strongly depend on wake onset because of the reduced threshold amplitude. However, I have not yet quantified how the solutions depend on prior conditions.

Dr. Kripke: One of the delights of the dual-oscillator model is that it explains the phase trapping that has been observed. Does your model display that?

Dr. Daan: I undertook one simulation. You saw a double plot in my presentation. It had occasional phase jumps.

Dr. Kripke: I saw phase jumps, but I did not see phase trapping. You can distinguish between the two, because phase trapping has an oscillation in phase without 360° jumps in phase.

Dr. Daan: Our model does not do that.

Dr. Weitzman: Your mathematical modeling of Dr. Borbély's concept provides a very powerful way to explain the relationship between a sleep-oscillator model and a sleep-deprivation model (or a relaxation oscillator), whatever the term you like to use. And if one uses Stage 3–4 as indicating the S (sleep) factor, and it is probably a very good approximation of that, one certainly finds that the longer one stays awake, there will be more Stage 4. After the long wake periods, there is more Stage 3–4 the first 2 hr than after the shorter wake periods, even though they do occur at different phases of the temperature cycle. However, there are a number of well-known phenomena that must also be explained. One is that long-term sleep deprivation leads to the intrusion of sleep processes, validated by EEG measurement, while the subject remains awake. That is, there are microsleeps, and there are changes in EEG waves. Second, if you sleep-deprive someone for a week, they are not going to sleep for 4 days straight. In the longest sleep-deprivation study ever scientifically studied, a subject stayed awake for 11 days straight and then slept for about 12 to 14 hr and felt perfectly fine thereafter. The explanation that I have is that his brain was sleeping, while he remained awake during those 11 days. Third, we know from many studies that the waking phenomenon will appear during sleep. For example, the presence of alpha waves will often occur in pathological and even natural conditions. We have seen that some of our subjects during free running show microsleeps. So that is another phenomenon in which the process of waking

occurred during the sleeping process. Lastly, your model does not take into account other kinds of rhythmic processes besides the sleeping–waking issue, namely, the timing of REM, the timing of temperature phase, the timing of hormonal (e.g., cortisol) phase. It seems to me that your one-oscillator model with an additional relaxation process is fine for sleep–wake phenomena, but then we have to move to some other control system, perhaps another oscillator system, for the other kinds of biological and physiological rhythms.

Dr. Daan: The thing that is characteristic of the model is a limit on the recovery process. However long you would keep your subject awake, you would never need more than 12 to 14 hr to recover. I cannot explain the other phenomena.

Dr. Borbély: Just to repeat these points about non-REM sleep, the idea is that REM reflects a circadian aspect of sleep, essentially governed by a sine wave. In addition, there is a slight sleep-dependent aspect of REM stages. These in conjunction explain most of REM sleep and explain the reciprocal interaction with non-REM sleep and REM sleep interacting negatively and inhibiting each other's peaks.

Dr. Weitzman: Well, I just have to say that Stage 3–4 sleep occurs during the first 2 hr after the subject goes to sleep, whether he be free running or entrained. The timing of REM sleep, however, shifts. It does *not* occur at the same time in free-running and entrained subjects. If REM sleep is timed, as you say, with respect to the sleeping process, and if your indicator of the sleeping process is Stage 3–4, it should occur in the same phase relationship whether you are entrained or free running.

Dr. Daan: No. The amplitude of the system is much lower during free running; so automatically it changes the phase relationship, and the process hits the threshold somewhere else. The phase differences between the two stages are thus easily explained.

Dr. McCarley: One of the most compelling pieces of evidence for a relatively independent ultradiar REM oscillator is the clinical phenomenon of narcolepsy, which represents a periodic breakthrough of REM sleep phenomena dissociated from slow-wave sleep phenomena with the same ultradian rhythm throughout the day, as it normally occurs only at night. I think that you have an excellent model because you are tying it to physiology. But I do not think it is necessary to have slow-wave sleep and REM sleep mutually inhibitory. They could instead be competing for expression. The neuronal or hormonal generators do not necessarily have to mutually inhibit one another.

Dr. Gander: In your simulations of free-running conditions, when you were getting very long sleeps and thus very long periods, did you see a progressive decline in the amplitude of the threshold?

Dr. Daan: Yes. However, it was very gradual.

Dr. Gander: I would like to make two points. You referred to our data on the monkey temperature rhythm, where there was a decline in amplitude between entrainment and free run. That could be explained by the masking effects of light; for example, in experiments with LD 2:2 cycles you can actually demonstrate that there is an increase in temperature during "lights on." Furthermore, the difference in amplitude between the entrained and free-running temperature rhythms can be explained purely in terms of the masking effects of light on the rhythm. The second thing is that the temperature rhythm might not be a good example to cite in your modeling, because it does not decline in amplitude in at least 140 days of free run in humans.

Dr. Daan: You maybe have a very good explanation for the free-running state, but I have referred only to the empirical evidence, that the amplitudes of circadian rhythms tend to be smaller in free run than in entrainment, whatever the cause of that may be.

Dr. Gander: I am just suggesting that that might not be an ideal example, because we believe we can explain the amplitude differences in another way.

Dr. Daan: The amplitude difference is not all that important. I have shown one simulation where you have a constant amplitude and you still get the initial free run.

Mathematical Models of the Circadian Sleep-Wake Cycle, edited by M. C. Moore-Ede and C. A. Czeisler. Raven Press, New York © 1984.

Looking at Human Circadian Phenomena from a Framework of Simple Stochastic Models

Gerhard Dirlich

Max Planck Institute for Psychiatry, 8000 Munich, West Germany

PRELIMINARY REMARKS

Self-Sustaining Circadian Rhythm: A Universal Phenomenon?

Research in circadian systems was initiated by the discovery of self-sustaining, approximately periodic rest–activity cycles under constant environmental conditions. Rest–activity cycles have been observed in some animal species under certain experimental conditions, and these are some of the temporally most precise biological rhythms known thus far (4, p. 16).

Such highly regular rest–activity patterns inspired the creation of the term *circadian clock*; moreover, they became the inductive basis of several mathematical models of circadian systems (4,10,13,14). These models are all aimed at explaining recurrent phenomena of high cycle-to-cycle precision; more randomly organized rhythms, which also exist, have to be treated by incorporation of sources of perturbation into the models. This is due to the intrinsic philosophy of these models, according to which circadian rhythms are basically periodic phenomena that under the influence of perturbing factors may appear more or less veiled or distorted (14, p. 99) (see R. A. Wever, *this volume*).

This implicit assumption may lead a layperson in this field of research to the conjecture that circadian fluctuations of life functions in general are distinct and temporally precise phenomena. No expert in the field, however, would deny the fact that this is not true. There are three important factors for the observation of circadian rhythms in constant environments, namely, the chosen species, the given experimental conditions, and the observed life function.

Persistent circadian rhythms can be observed only for certain configurations of these factors; i.e., for a given species, certain species-specific environmental conditions and an appropriate choice of the observed life function are required (5, p. 14).

For example, it has been shown experimentally that rest–activity behavior is not always organized in the form of circadian cycles (8). Moreover, it has been explored systematically how self-sustainment of precise circadian rhythms

depends on certain environmental variables. For instance, rest–activity cycles persist only when the light intensity lies between critical lower and upper thresholds. With values of the intensity outside this interval, the animals develop more or less irregular rest–activity patterns.

The existence of regular 24-hr fluctuations has been demonstrated in many biological and behavioral variables (13, p. 3). However, despite the vast literature on experiments under constant conditions, our knowledge about circadian rhythms under these circumstances is still limited with respect to the entire spectrum of candidate variables: the majority of investigations in animals have been concerned with the rest–activity behavior; a much smaller amount of data on body temperature is available, and there are relatively few data on endocrine variables. In humans, the observational data comprise the rest–activity cycle, the body temperature, some parameters of the blood-circulating system, some ingredients of blood plasma and urine, and finally some psychometric variables.

Thus, circadian rhythms as distinct and temporally precise phenomena are not universal, but rather an indication of some special mode of operation of the biological timing mechanisms in cases of favorable configurations of species, environment, and observed variable. For certain other configurations of these factors it has been demonstrated that the circadian rhythms degenerate in constant environments.

This delimits the validity of oscillator models of circadian systems. They are suitable only for approximately periodic phenomena; they do not allow for the description of aperiodically recurring events.

In this situation it appears attractive to consider the possibility of developing a more general biological theory of the temporal organization of life functions, whose range of validity includes the case of ongoing distinct and precise circadian rhythms as a special case. Some elements of a theory of this kind would certainly have to be borrowed from mathematics, just like the concept of oscillators, which was borrowed from physics. This chapter is concerned with a class of mathematical models that might be useful for such an approach. They are based on the concept of stochastic processes. This type of model is suitable for the description of irregular, random phenomena, as well as for almost perfectly periodic rhythms. Thus, in contrast to oscillator models of the circadian clock, periodicity is here not an assumption but a property resulting from a special choice of model parameters or the special structure of the model.

Overview

This chapter consists of three parts. In the first, the basic philosophy of stochastic models is outlined, and a simple type of stochastic model—the renewal model—and its performance are described in detail. The second part deals with an application of this model to human rest–activity data. Empirical data are analyzed with respect to the question of how precise the human

circadian clock is in the state of internal synchronization. The last part reports an application of the renewal model in the framework of a complex system of several simultaneously operating processes, describes some features of its performance, and discusses its relevance as a model of a candidate mechanism for the generation of circadian rhythms.

STOCHASTIC MODELS

Although an introduction of a mathematically correct notion of stochastic models (6) is beyond the scope of this chapter, an outline of the basic philosophy and possible applications may help to explain the following paragraphs. Let us first consider a simple physical system whose dynamics can be described by a stochastic model.

A Storing System

The system is a container with an input and an output channel. It operates in the following way: Initially the container is empty, and the output channel is closed. Then, through the input channel, the container is filled. When it is full, the input is shut off, and the output channel is opened. The container's content is released. As soon as the container is empty, the cycle begins anew (Fig. 1). Each cycle consists of an alternation of the two states of operation. Let the amounts of input and output per time unit be randomly varying quantities. Then, obviously the time intervals required for filling and emptying the container (i.e.,

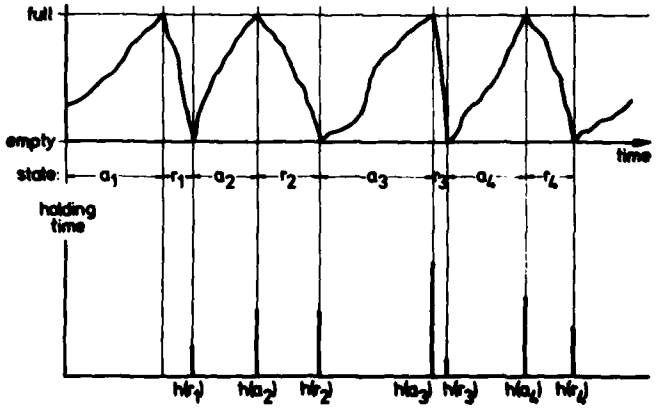


FIG. 1. Storing system with two states of operation (*a*, filling; *r*, emptying). **Top:** Time course of the instantaneous content. **Bottom:** Sequence of the holding times of the alternating states.

the holding times of the two states of operation), as well as the cycle duration, are stochastic quantities.

Assume that we have been watching this system for some time through a series of several cycles. Although we know part of its history exactly, in principle we are unable to predict the future holding times with certainty. This is because of the stochastic nature of input and output. The important feature of this system is the existence of many possible continuations of the process at each time. This is the common property of processes for which stochastic models are suitable.

The explicit formulation of a stochastic model requires additional specific knowledge to be expressed in the form of model assumptions. In the following, a model for the system described here is constructed, starting from a set of intuitively appealing assumptions.

Systems with Two States

Before going into details, let us reconsider the general case of cyclic processes with two alternating states. Formally similar to the system discussed here, the rest–activity cycle can be regarded as a cyclic process with two states, one representing activity and the other representing rest. Besides the correspondence of their state spaces, the previously described storing system and the rest–activity cycle can be related in a more substantial respect. The stored content may correspond to the hypothetical sleep factor that is assumed to accumulate during activity and to decline during sleep (1) (see S. Daan and D. Beersma, *this volume*). This correspondence, however, is not further discussed here, because this chapter is mainly concerned with model-related considerations of cyclic processes.

A cyclic process with two states may be characterized by the state diagram shown in Fig. 2. Within each time unit, exactly one state transition occurs. During uninterrupted sequences of transitions of type 1 or 4, the system stays in state a or r, respectively, for certain holding times, labeled $h(a_i)$ and $h(r_i)$ in the following. A realization of the process can be described by a vector of holding times:

$$\langle h(a_1), h(r_1), h(a_2), h(r_2), \ldots, h(a_n), h(r_n) \rangle$$

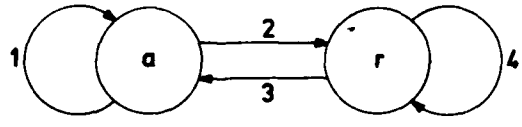


FIG. 2. State diagram of a process with two states labeled *a* and *r*. Transitions are labeled 1 through 4.

The Alternating Renewal Model

Now, let us introduce a simple probability model for this process. Its probability assumptions are the following:

1. $\langle H_{a_1}, H_{r_1}, \ldots, H_{r_n} \rangle$ is a series of random variables, and $\langle h(a_1), h(r_1), \ldots, h(r_n) \rangle$ is a sample.
2. All H_{a_i} $(i = 1,n)$ have the same probability distribution function F_a (i.e., they are identically distributed), and analogously the H_{r_i} $(i = 1,n)$ are identically distributed with F_r.
3. All pairs (H_x, H_y), with x and y elements of $\{a_i,\ i = 1,n\} \cup \{r_i,\ i = 1,n\}$ and $x \neq y$, are stochastically independent. These assumptions define the stochastic model of an alternating renewal process (6, p. 278).

Originally, the stochastic renewal model was suggested for applications to physical systems with components that expired after a variable time of operation and had to be replaced by new ones. The renewal paradigm has been used in models of circadian clocks (4), in which cyclic behavior of neurons is depicted by alternating renewal processes. Another biological application was suggested by Lehmann (7). He has used alternating renewal models to describe irregular rest–activity patterns.

These hints may suffice to demonstrate the range of possible applications of the stochastic renewal model. Although the performance of the model is well known, it appears useful to describe here some basic features. Its performance is determined by the assumptions stated earlier and by the probability distri-

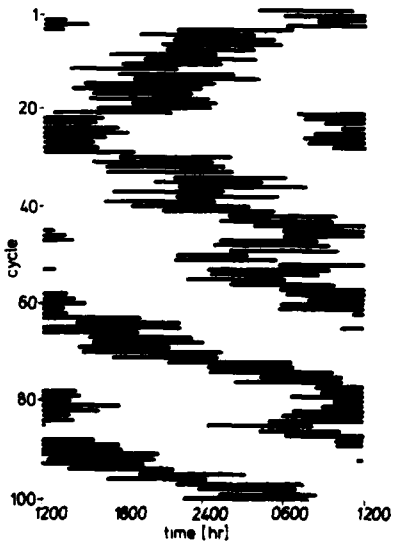


FIG. 3. Chronobiological standard plot of a series of 100 cycles of an alternating renewal process. State *a*, *bars*; $\mu_a = 16.2$ hr; $\sigma_a = 3.0$ hr; $\mu_r = 8.0$ hr; $\sigma_r = 1.0$ hr.

butions F_a and F_r. In the following, it is assumed that F_a and F_r are parametric probability distributions: i.e., Gaussian distributions that are determined by the parameters μ_a and σ_a and μ_r and σ_r, respectively.[1] Although the performance of the model is, in principle, qualitatively the same for any choice of parameter values, it looks very different when depicted in the form of the well-known chronobiological standard plot (Fig. 3).

In the alternating renewal model, simple laws hold for the expectation and the standard deviation of sums of holding times, especially for the sum of an activity-phase holding time $h(a_i)$ and the following rest-phase holding time $h(r_i)$, which constitute the cycle duration τ_i:

$$\mu_\tau = \mu_a + \mu_r; \qquad \sigma_\tau = \sqrt{\sigma_a^2 + \sigma_r^2}$$

In the case presented here, the expectation of the cycle duration is $\mu_\tau = 24.2$ hr, and the standard deviation of the cycle duration is $\sigma_\tau = 3.16$ hr. The quantity v_τ defined as σ_τ/μ_τ is a parameter that characterizes the relative cycle-to-cycle imprecision of the process. In this case, $v_\tau = 0.130$.

From Fig. 3 it is obvious that it is not possible in this case to make predictions about the future of the process even after a long time of observation: Between cycles 5 and 17 the transition times drift to the left; then until cycle 30 they oscillate around a fixed time; then a phase jump of about 6 hr occurs during cycle 31; after cycle 64, the transition times drift to the right. Note that all these events occur at random; the system has been stationary during the entire interval of 100 cycles.

Figure 4 shows a realization of the alternating renewal model with reduced values of the standard deviation of the holding times. Here, $\sigma_\tau = 0.71$ hr, and $v_\tau = 0.029$. Now the pattern looks much more regular. However, also in this case, an estimation of the period from sequences of 20 to 30 cycles would yield misleading results. Take, for instance, cycles 20 to 42 and 73 to 100. We would estimate $\bar{\tau}$ values of 24.42 hr and 23.86 hr. It should be emphasized here that an important feature of the performance of the alternating renewal model is the slow fluctuations and drifts of the transition times along the time axis that make it impossible to predict the phase of these events with respect to the geophysical time for cycles that are more than a few days ahead.

When the cycle-to-cycle precision is further increased, the performance of the alternating renewal model resembles more and more a periodic process. This is shown in Fig. 5. The imprecision here is $v_\tau = 0.006$.

These examples reveal some aspects of the performance of the alternating

[1] It is mathematically not correct to use Gaussian distributions for holding times, because the probability of negative holding times is theoretically not zero. However, for appropriate choices of μ and σ, this probability can be neglected in practice. Therefore, Gaussian distributions are used in the following.

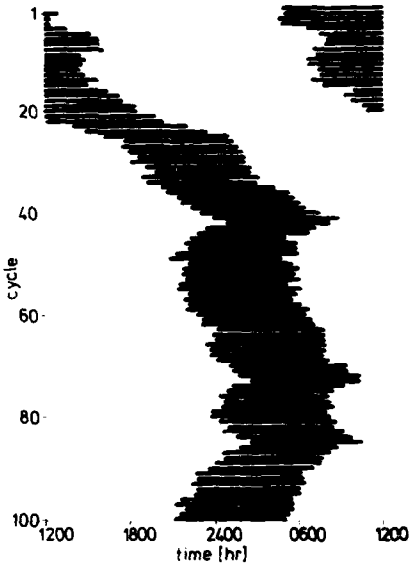


FIG. 4. Series of 100 cycles of an alternating renewal process; $\mu_a = 16.2$ hr; $\sigma_a = 0.5$ hr; $\mu_r = 8.0$; $\sigma_r = 0.5$ hr.

renewal model. Their purpose is to give an impression of how the temporal state pattern represented in the form of a chronobiological standard plot is affected by the variance of the holding times of the two states. Obviously, it is possible to generate approximately periodic patterns, as well as highly irregular patterns,

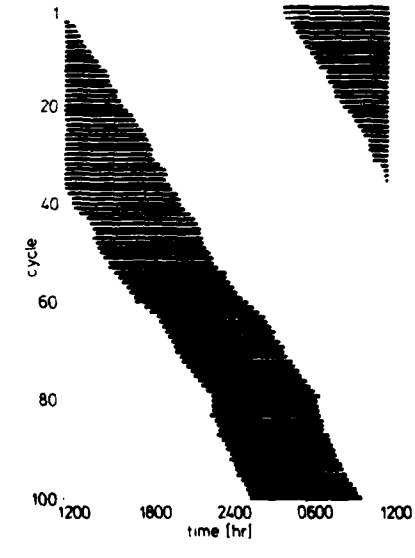


FIG. 5. Series of 100 cycles of an alternating renewal process; $\mu_a = 16.2$ hr; $\sigma_a = 0.1$ hr; $\mu_r = 8.0$ hr; $\sigma_r = 0.1$ hr.

by choosing appropriate values of the variances of the holding times. For mathematical investigation, the model can also be used as a statistical basis for an analysis of observational data by testing its assumptions and estimating its parameters. Such an investigation may be regarded as part of a search for appropriate models that can account for irregular rest–activity data as well as for rest–activity cycles with higher cycle-to-cycle precision.

IMPRECISION OF THE HUMAN CIRCADIAN CLOCK IN THE STATE OF INTERNAL SYNCHRONIZATION

This section reports an empirical investigation of the cycle-to-cycle precision of human rest–activity data observed during temporal isolation of the subject. It is based on a mathematical model of a clock and an overt rhythm that is coupled to it.

How Precise Is the Circadian Clock in Humans in the State of Internal Synchronization?

A visual comparison of standard plots of human rest–activity data from experiments in temporal isolation (2,13,15) (see R. E. Kronauer, *this volume*) with plots from animal experiments under constant conditions (11) shows that the cycle-to-cycle precision of the human rest–activity cycle in most cases is significantly lower. In fact, almost randomly organized rest–activity patterns have been observed in humans (3). Let us consider this in connection with two well-known statements about circadian clocks and the human rest–activity cycle: First, it is commonly assumed that the precision of circadian clocks increases with increasing complexity of the organism. Second, the human rest–activity cycle is regarded as a relatively unreliable indicator of the clock cycle. Thus, the question arises whether or not the obviously greater imprecision of the human rest–activity cycle can be fully attributed to a weak interconnection between a precise underlying clock and the overt process. If this is not the case, the circadian clock must be assumed to generate a cycle with variable period. Supposedly, this may be because of the given experimental conditions: Even in time-cue-free environments, some known and perhaps also some still unknown factors may affect the circadian clock, causing instantaneous phase shifts that vary its cycle duration. In light of this, another question arises: Is it possible at all to create environmental conditions that favor the sustainment of precise circadian rhythms in humans? If such conditions cannot be established, the validity of models implying a precise clock will be severely restricted. These considerations triggered the investigation reported here; its objective was to estimate the imprecision of the human circadian clock in the state of internal synchronization.

TABLE 1. *Empirical data*

Source (Ref.)	Identification of data set[a]
Czeisler (2)	FR01, FR02, FR05, FR06, FR07, FR09, FR131 8 9 13 13 14 14 28
Kronauer (*this volume*)	02 32
Lund (9)	154, 158, 160, 164, 167, 169 14 13 12 13 12 12
Wever (13)	HS61, HS62, AG, DB21, MS23, AS, DP, HZ 27 22 31 24 23 27 28 17
Zulley (15)	CMoZ, CMmZ, MCmZ, MCoZ, AFmZ, WLmZ, WLoZ, MKmZ 22 29 29 19 19 13 13 11

[a]Identification labels are those used in the sources. The number of cycles included in the analysis described here are indicated below the labels.

Empirical Data

The data analyzed in this investigation come from different sources (Table 1). During the time intervals analyzed here, all subjects were in the state of internal synchronization. Figure 6 shows the distribution of the standard deviations of cycle durations s_t in these subjects.[2] The standard deviations lie in the range from 0.4 hr to 3.4 hr. The imprecision v_t varies in the range from 0.017 to 0.141. These values cannot directly be interpreted as estimates of the precision of the clock cycle, because the overt rest–activity cycle is only a fuzzy indicator of the assumed underlying clock cycle. However, estimates of the clock cycle precision can be obtained by analyzing the data on the basis of a model of a clock and an overt cyclic process coupled to it.

Model of a Clock and an Overt Rhythm Coupled to It

Pittendrigh and Daan (12) outlined such a model. It is schematically represented in Fig. 7. They described a method to split the variance of the rest–activity cycle duration s_t^2 into two components, one of which gives an estimate of the variability of the oscillator-cycle duration s_τ^2 and one of which gives an estimate of the coupling strength between the oscillator and the overt rhythm s_w^2. In this approach, the serial correlation r_s between the durations t_i and t_{i+1} of successive cycles is of central importance.[3]

[2]Cycles were defined from the beginning of an activity phase to the beginning of the next activity phase.

[3]The variance of the clock-cycle duration is estimated by $s_\tau^2 = (1 + 2r_s)s_t^2$, and the strength of the coupling by $s_w^2 = -r_s s_t^2$ (12).

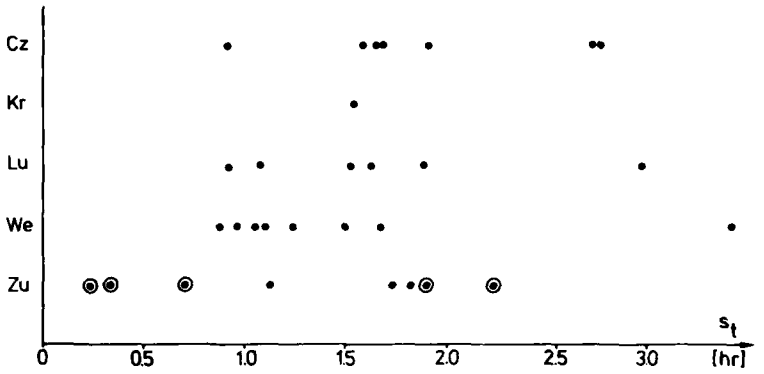


FIG. 6. Frequency distribution of the standard deviation of the duration of the rest–activity cycle s_t. The labels at the *y* axis identify the different sources of the data (see Table 1). Circles indicate experiments with Zeitgeber.

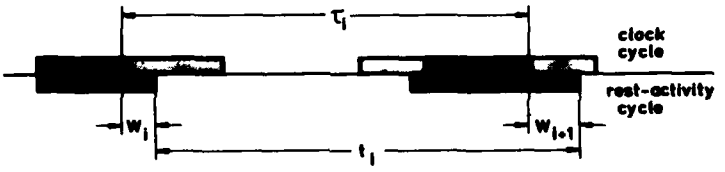


FIG. 7. Schema of a model of a clock and an overt rhythm (rest–activity cycle) coupled to it; τ_i is duration of the clock cycle; t_i is duration of the rest–activity cycle. Shaded areas are time intervals during which state transitions of the rest–activity cycle must occur; w_i is the difference of transition times in clock cycle and overt cycle.

Besides this method, there are other computational procedures that allow for separation of the previously mentioned sources of variability. Estimates of s_τ^2 can be derived from s_{2t}^2 and s_{3t}^2, the respective variances of the series of sums of two and three adjacent cycle durations[4]:

$$s_\tau^2 = s_{2t}^2 - s_t^2$$

$$s_\tau^2 = (s_{3t}^2 - s_t^2)/2$$

[4]Besides $\sigma_t^2 = \sigma_\tau^2 + 2\sigma_w^2$, also $\sigma_{2t}^2 = 2\sigma_\tau^2 + 2\sigma_w^2$ holds. Thus, $\sigma_\tau^2 = \sigma_{2t}^2 - \sigma_t^2$.

In the investigation reported here, the formula of Pittendrigh and Daan as well as these formulas were used to estimate the respective parameters of the model. In addition to the variance of the cycle duration s_t^2, also s_τ^2, r_s, and s_w^2 were estimated. Besides an analysis of the observational data listed in Table 1, the statistical properties of these parameters were investigated by Monte Carlo studies.

Results

Serial Correlation Coefficient

The first result concerns the serial correlation coefficient r_s that has been interpreted as a discriminator between cycles of oscillatory and stochastic origin (13, p. 32). Approximations of the probability distribution of r_s in dependence of the model parameters σ_τ and σ_w were computed by Monte Carlo techniques.

For different choices of the model parameters σ_τ and σ_w, realizations covering 14, 21, and 28 days were computed. From these data, the distributions of s_t, r_s, s_τ, and s_w were estimated. Figure 8 shows the regions containing 95% of the r_s values in the case of realizations over 14 days. For the four choices of σ_τ and σ_w represented here,[5] the tolerance intervals overlap each other to a large extent. Therefore, it is not possible to infer the imprecision of the underlying clock solely from the serial correlation coefficient. Pittendrigh and Daan's

[5]In the cases $\sigma_\tau = 0$ and $\sigma_w = 0$, distributions with very small variances ($\sigma^2 = 0.0001$) were used.

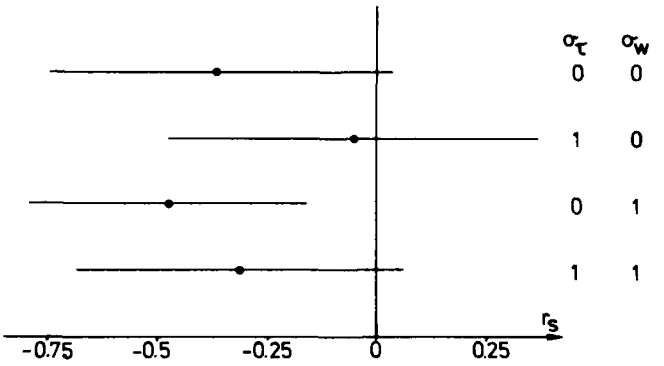


FIG. 8. Ninety-five percent tolerance intervals of the serial correlation coefficient r_s (*left*). From top to bottom: Intervals for four choices of the model parameters (*right*) computed by Monte Carlo techniques. Tolerance intervals are represented by *bars*, mean values by *dots*.

model shows that a joint analysis of the variance of the overt cycle s_t^2 and the correlation r_s is necessary.

The values of r_s computed from the observational data are represented in Fig. 9. The median of their distribution is -0.13. This value lies outside the tolerance interval only for the case $\sigma_\tau = 0.0$ hr and $\sigma_w = 1.0$ hr. These results are a by-product of the investigation. We shall now return to the main line of our considerations, which are aimed at achieving an estimation of the imprecision of the human circadian clock.

Imprecision of the Clock Cycle

Estimates of the imprecision of the clock cycle from single series of rest–activity cycles are unreliable, because s_t, s_τ, and s_w also have overlapping distributions, similar to the situation just demonstrated concerning r_s. Therefore, the results presented in the following are interpreted only from a group-statistical point of view.

In Fig. 10, the values of s_τ are displayed, together with the upper limit of the 95% tolerance interval of s_τ for the case in which the model parameters are $\sigma_\tau = 0.2$ hr and $\sigma_w = 1.0$ hr. For 11 of 30 subjects (37%), the hypothesis that the clock cycle has a high precision ($s_\tau < 0.2$ hr), i.e., $v_\tau < 0.009$, must be rejected. On the other hand, all of these data are compatible with the hypothesis of an imprecise clock cycle ($s_\tau > 1.0$ hr), i.e., $v_\tau > 0.04$.

Strength of the Coupling

Finally, the results concerning the strength of the coupling between the clock and the overt cycle are briefly reported. The strength of the coupling between

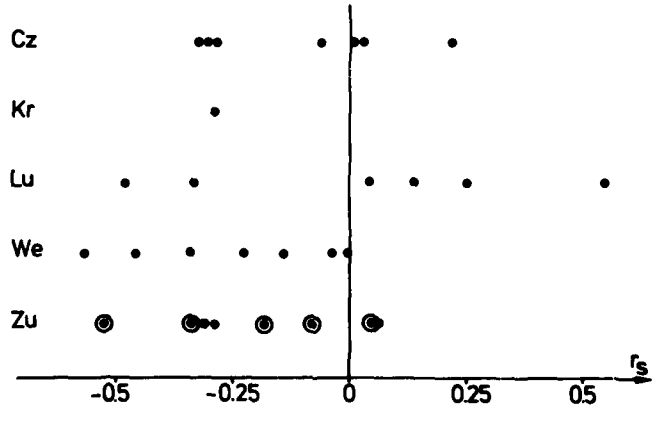


FIG. 9. Frequency distribution of the serial correlation coefficient r_s (see Fig. 6).

model shows that a joint analysis of the variance of the overt cycle s_t^2 and the correlation r_s is necessary.

The values of r_s computed from the observational data are represented in Fig. 9. The median of their distribution is -0.13. This value lies outside the tolerance interval only for the case $\sigma_\tau = 0.0$ hr and $\sigma_w = 1.0$ hr. These results are a by-product of the investigation. We shall now return to the main line of our considerations, which are aimed at achieving an estimation of the imprecision of the human circadian clock.

Imprecision of the Clock Cycle

Estimates of the imprecision of the clock cycle from single series of rest–activity cycles are unreliable, because s_t, s_τ, and s_w also have overlapping distributions, similar to the situation just demonstrated concerning r_s. Therefore, the results presented in the following are interpreted only from a group-statistical point of view.

In Fig. 10, the values of s_τ are displayed, together with the upper limit of the 95% tolerance interval of s_τ for the case in which the model parameters are $\sigma_\tau = 0.2$ hr and $\sigma_w = 1.0$ hr. For 11 of 30 subjects (37%), the hypothesis that the clock cycle has a high precision ($s_\tau < 0.2$ hr), i.e., $v_\tau < 0.009$, must be rejected. On the other hand, all of these data are compatible with the hypothesis of an imprecise clock cycle ($s_\tau > 1.0$ hr), i.e., $v_\tau > 0.04$.

Strength of the Coupling

Finally, the results concerning the strength of the coupling between the clock and the overt cycle are briefly reported. The strength of the coupling between

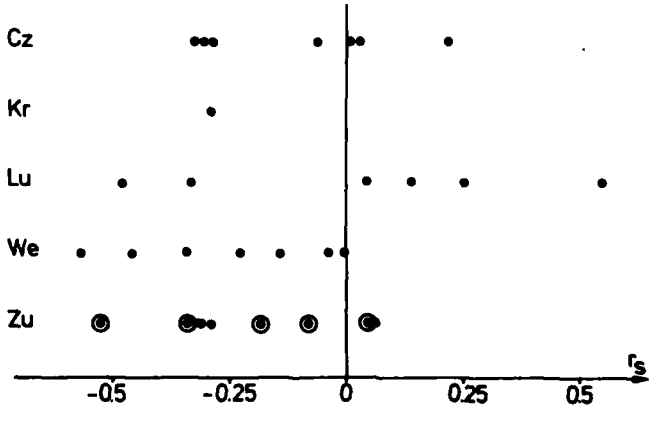


FIG. 9. Frequency distribution of the serial correlation coefficient r_s (see Fig. 6).

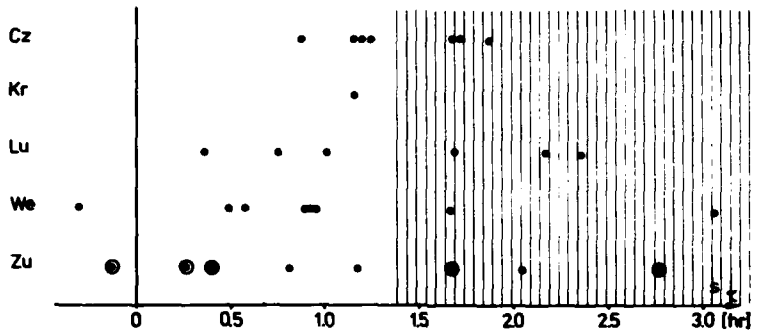


FIG. 10. Frequency distribution of the standard deviation of the duration of the clock cycle s_τ (see Fig. 6). Shaded area: this interval under the hypothesis $\sigma_\tau < 0.2$ hr and $\sigma_w = 1.0$ hr has a probability of less than 5%.

the clock and the overt rhythm determines the variance s_w^2. In Fig. 11, the distribution of the estimated s_w values is shown. The median of the distribution is 0.6 hr. This is 38% of the median of s_t, which is 1.6 hr. These figures show the contribution of the stochastic coupling between the clock and the overt cycle to the variance of the duration of the overt cycle.

Discussion

The results reported here show an average imprecision of 7% for the human rest–activity cycle. The most precise circadian rhythms in animals observed thus far have an imprecision of about 0.2%. These figures emphasize the difference in the cycle-to-cycle precision of circadian rhythms in temporally isolated humans and in animals under certain constant conditions. The estimated imprecision of the cycle of the human circadian clock is significantly smaller; it has a magnitude of 4% (median of the σ_τ distribution: 1.0 hr). Even for this value of s_τ, however, the 95% confidence interval for the duration of the clock cycle covers about 4 hr. In other words, the data analyzed here indicate a considerable cycle-to-cycle variability of the human circadian clock even in the state of internal synchronization.

However, the model applied here does not provide a fully satisfactory explanation of this result, in that not only random variability of the duration of the clock cycle contributes to s_τ but also systematic variations of the period of the clock. Examples of types of such systematic variations are the following: trends of τ_i phase shifts of the rest–activity cycle that are reversed after several days ("weekend shift"); variations of the period in cases of relative coordination.

The model used here does not allow for a separation of factors that cause nonrandom (i.e., systematic) variations of the cycle duration from random

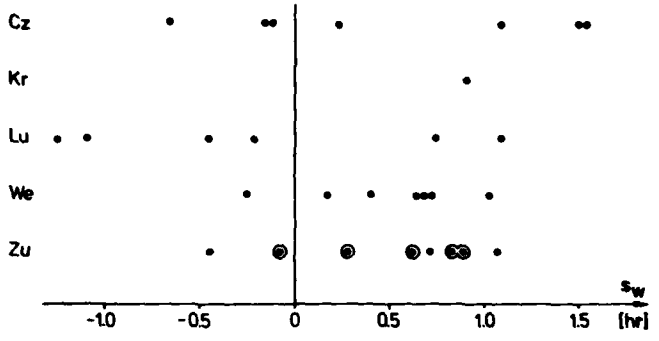


FIG. 11. Frequency distribution of the standard deviation s_w characterizing the strength of the coupling (see Fig. 6).

effects. The observations analyzed in this investigation, however, were all obtained from subjects in a steady state of internal synchronization.[6] Thus, the previously mentioned major sources of systematic variability were not in operation, and we must acknowledge that the period of the human circadian clock varies randomly in the state of internal synchronization. This may be an intrinsic property of the clock, or it may be the result of randomly induced phase shifts in the clock cycle.

The first hypothesis according to which the variability of the period is a property of the clock is compatible with the concept that the entire set of temporally coordinated cyclic life functions performs cycles whose periods vary randomly due to random fluctuations in the cycle durations of the functions participating in this system. The second hypothesis implies the existence of a clock-like organ that is (in contrast to a physical clock) not completely shielded from external influences and is reset by certain factors on a random schedule (C. Eastman, *this volume*). In the last part of this chapter, the first of these hypotheses is illustrated by the conception of a mathematical model, whose performance is then analyzed.

A NETWORK OF RANDOM PROCESSES

In the following, a system consisting of several simultaneously operating random processes, a network of random processes (NORP), is described. These processes are interconnected by a network of directed relations. It is shown that

[6]This state is essentially different from the transient states of beginning desynchronization that have been analyzed empirically (2,13) as well as through the study of mathematical models (R. A. Wever, *this volume*, and R. E. Kronauer, *this volume*).

the system is capable of sustaining relatively precise cycles, with all participating processes staying mutually synchronized. Several aspects of the performance of the system are demonstrated, and implications of the model are outlined.

Biological Frame of Reference

Some theories of the circadian clock assume that the observable physiological and behavioral variables represent processes external to but functionally connected with the clock (*vide supra*). According to these theories, the clock controls the overt circadian rhythms (13, p. 236). From an alternative point of view, the entire variety of cyclic physiological and behavioral functions of the organism may be regarded as constituents of the circadian clock. According to this concept, the circadian clock is a certain mode of operation of the set of circadian and more rapidly cycling life functions. An important component of a model based on this philosophy is the assumption of mechanisms that enable mutual control of the timing of the participating processes. Would a system equipped with such mechanisms be able to exhibit self-sustaining cycles, with the participating processes staying internally synchronized in constant conditions? What are the minimal requirements concerning the structural details of the control mechanisms and the structure of the network of interconnections to enable the system to generate self-sustaining cycles? This section deals with questions of this kind.

Some features of the model envisioned here formally resemble Enright's coupled stochastic system (4). There are, however, also important differences. Enright's model maps onto a different part of the biological reality, i.e., onto the structure and performance of neuronal networks, whereas the model conceived here is aimed at depicting another layer of organismic functions, which in comparison with processes in single cells and delimited neuronal networks are rather global variables characterizing the state of the organism. Examples include body temperature, the sensation of hunger, and the excretion rate of adrenalin. Some of these processes are consciously perceivable, and others are not. This layer of life functions with its intrinsic structure of interconnections is regarded here as a system that may essentially contribute to the self-sustainment of circadian rhythms in temporal isolation. Especially, some phenomena observed in temporal isolation of humans, such as the state of internal desynchronization, may be the result of processes in this layer of body functions.

Our present knowledge about the functional dependencies between different physiological systems and about the interaction between physiological and cognitive processes in humans is, of course, insufficient for the conception of a structural model for this complex system. Nevertheless, even a speculative model whose purpose may be an exploration of the structure and performance

of a certain type of complex system may contribute to the generation of new ideas and to the development of new theories of circadian rhythms.

The following consideration played important roles in the conception of the model: In an organism that has lived for a long time in an environment with a periodically recurring 24-hr schedule of events and tasks, many body functions have become temporally coordinated with each other in the form of a timetable that may be stored in memory. When the organism is exposed to time-cue-free conditions, two antagonistic forces start to operate. On one hand, each physiological process and each behavioral cycle has a tendency to develop its own free-running schedule. On the other hand, the organism as a whole may have a preference for maintaining the normal learned pattern of temporal coordination of its subsystems to which it is accustomed.

The Orchestra Metaphor

The main features of the system suggested here can be vividly described by comparison of a biological system with an orchestra. The system corresponds to the entire orchestra; each participating process corresponds to one of the musicians. Each musician has on his music stand a sheet containing his part of the piece of music to be played. The notes constitute a memory from which he can read what he is supposed to play. The conductor has a score on his music stand containing the parts of all musicians. He takes care of the timing of the orchestra. Let us assume that at some time the conductor leaves and the orchestra continues playing without a pacemaker, and each musician goes on playing his part without paying attention to the playing of his colleagues. Sooner or later the musicians will run out of synchrony and will become mutually desynchronized with each other, because each musician's part consists of a series of tones whose holding times are subject to small random variations. The performance of each musician in this state can appropriately be described by a renewal model with several states, each corresponding to a note.

Let us slightly change the scene (Fig. 12). Each musician receives a sheet that shows not only his own part but also the part of at least one other musician in the orchestra. Now, even if the conductor leaves, each musician can listen to the play of those colleagues whose parts he can compare with his own part as a reference for his timing. He will notice if he is relatively behind or ahead of the schedule and will correct his play, if necessary. Several relations of this kind constitute a network interconnecting the musicians. This metaphor demonstrates the essential features of the mathematical model suggested here.

A Realization of NORP

The model was realized in the form of a computer program. This kind of realization has several obvious advantages over the classic form of representing mathematical models by closed formulas (R. A. Wever, *this volume* and R. E.

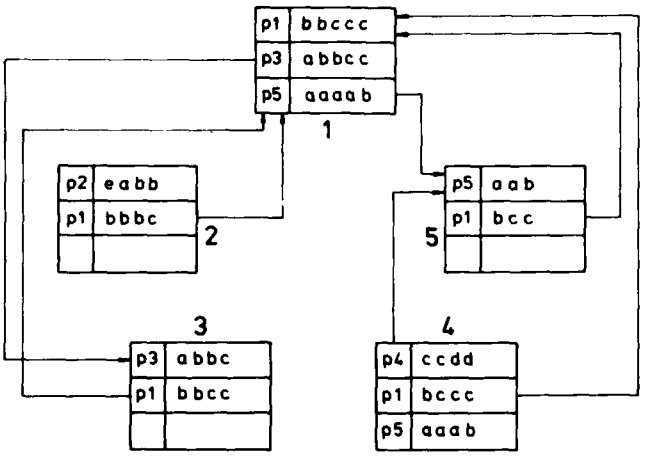


FIG. 12. A NORP with five processes. Each process has a memory (box) containing information on the temporal coordination between its scheduled sequence of states (row 1) and the scheduled sequences of states of one or several other processes (consecutive rows). Each state transition is represented by a state vector (column) in the memory. Arrows indicate directed relations between the processes.

Kronauer, *this volume*). First, a computer realization makes the structure of the model transparent, so that missing elements and inconsistencies of the structure can be detected. Second, a computer model is easy to modify. This allows for comparisons of different versions of the model. Third, the performance of the model can be explored by Monte Carlo techniques.

Structure

The computer realization of the model described in the following is a very simple and special version of NORP. An understanding of its performance, however, requires some details about its structure that will be outlined next.

Cyclic Random Processes

The implemented NORP is constituted by 10 cyclic random processes $P_1, \ldots, P_{10}$. Each process cycles through 10 states $x = a,b, \ldots, i,k$. The holding times in all states of all processes are identically distributed according to a Gaussian distribution with parameters μ_x and σ_x. When the processes run independent of each other, they are renewal processes, in essence. Thus, they have the following property: The expectation of the cycle duration is $\mu_\tau = 10\mu_x$, and the standard deviation of the cycle duration becomes $\sigma_\tau = (10)^{1/2}\sigma_x =$

$3.16\sigma_x$.[7] The instantaneous state of the system at time t is described by the state vector

$$\langle x_1(t), x_2(t), \ldots, x_{10}(t) \rangle$$

with $x_k(t)$ symbolizing the state of process P_k at time t (Fig. 13, bottom). If $\sigma_x > 0$, the probability for a given state vector to recur periodically is very small. The system described thus far is analogous to an orchestra without a conductor and with each musician having only his own part on his music stand.

[7] In the general case of n processes with states $x_1, \ldots, x_{mk}$ for process P_k, the holding time in state x_l of P_k is given by a Gaussian random variable with μ_{lk} and σ_{lk}. If the holding times are independent of each other, the distribution of the cycle duration τ_k or process P_k has the following parameters:

$$\mu_{\tau_k} = \sum_{l=1}^{m_k} \mu_{lk}; \qquad \mu_{\tau_k} = \left(\sum_{l=1}^{m_k} \sigma_{lk}^2 \right)^{1/2}$$

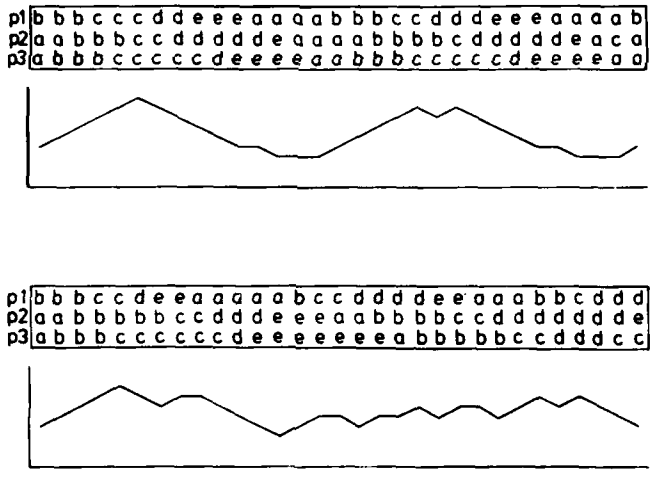


FIG. 13. Pattern of temporal coordination between three processes. **Top:** State of internal synchronization. **Bottom:** State of asynchrony. Each process cycles through five states (*a–e*). During internal synchronization, a cycle is terminated after 15 transitions (15 state vectors). Functions of the state vectors may be defined. The curves represent the course of a function of the state vector with clear periodicity in the state of internal synchronizaton and random fluctuations in the case of asynchrony.

Directed Relations: The Elements of the Network

The modification of the orchestra described earlier is now performed accordingly with our model. Components are incorporated into the system that enable it to maintain a predetermined pattern of temporal coordination between the different processes. The basic idea is that this control mechanism is solely based on directed relations between pairs of processes; i.e., it is of the form $P_k \rightarrow P_{k'}$. In other words, a central component that is connected to all P values is not assumed. This feature constitutes an important structural difference in comparison with the coupled stochastic system (4). The control mechanism suggested here is decentralized.

How does $P_k \rightarrow P_{k'}$ work? If a distinct pattern of temporal coordination between the two processes recurs repeatedly (Fig. 13, top), a simultaneous observation of the processes $(P_k, P_{k'})$ can be represented in the matrix symbolizing all possible states of $(P_k, P_{k'})$. Let $k = 1$ and $k' = 2$. Imagine that P_1 and P_2 operate simultaneously (Fig. 14). At some time, P_1 will change its state; some time later, either P_2 or once again P_1 will change its state. Let us assume that P_1 is presently in state a, and P_2 is in state d. The next transition may be made by P_1, which goes into state b. Then it assumes the states c and d, while P_2 is still in state d. Now, P_2 makes a transition to state e; then P_1 goes to

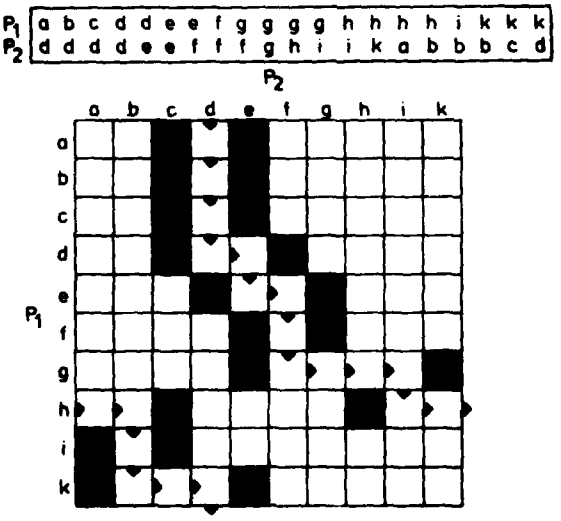


FIG. 14. Joint state space of two processes. A predetermined cyclic pattern of temporal coordination is indicated (*arrows*). The shaded fields represent states in which control mechanisms are activated.

e, and so on. Such a series of state transitions describes the pattern of temporal relationship between processes 1 and 2. Let us assume that after some more transitions the system returns to the initial state (a, d).

In case process 1 is related to process 2, process 1 has a memory containing a matrix with a cyclic path of state transitions representing the schedule of the temporal coordination between the two processes. Because both P_1 and P_2 are random processes, in the sense that the actual holding times in each of their states are random quantities, it may happen that the system leaves the path that is intended by the schedule. Consider, for instance, the following situation: When the system is in state (d, d), it may occur that the next transition is made by P_1. In this case, the system will deviate from the path and enter state (e, d). Now, a mechanism is activated that slightly increases the probability that the system may return to the memorized path. In the present case, an appropriate action will be to slow down P_1 by increasing the expectation of the holding time of its present state. This gives P_2 a greater chance to make a transition to state e while P_1 is still in state e. Thus, it is not unlikely that the system may reach the memorized path in state (e, e) again. This mechanism is activated whenever P_1 notices that the system is in one of the shaded fields adjacent to the path.

These details may suffice to show how directed relations between two processes are realized that enable the "looking" process to augment the chance to correct small deviations from its predetermined schedule of temporal coordination. Two important features of the model must be emphasized here: It does not contain a central unit to which all participating processes are connected, and there are no devices in the model that measure time, but only mechanisms that attract a process to the path in case of a deviation.

Types of Networks

In a system with 10 processes, a variable number of relations of the kind described here may exist: On one end of the spectrum there is the case of mutual independence (i.e., there are no relations at all); on the other end, when each process is related to each of the remaining 9 processes, there are 90 relations.

Different types of organization of the network of relations are conceivable (Fig. 15). Relation R1 represents the case in which each process is related to each remaining process. In case R2, one process is the pacemaker to which all the other processes are related, and R3 represents a hierarchical network of the existing relations. In R4 the relations form a cycle connecting all processes, and R5 represents a system with two independent networks. Relation R0 represents the case in which there are no relations between the processes.

The features of the NORP version reported thus far may suffice to explain the concept of the mathematical model suggested here. Moreover, they provide a frame of reference for some aspects of the performance of the model that will be discussed next. It should be noted that the model has great flexibility.

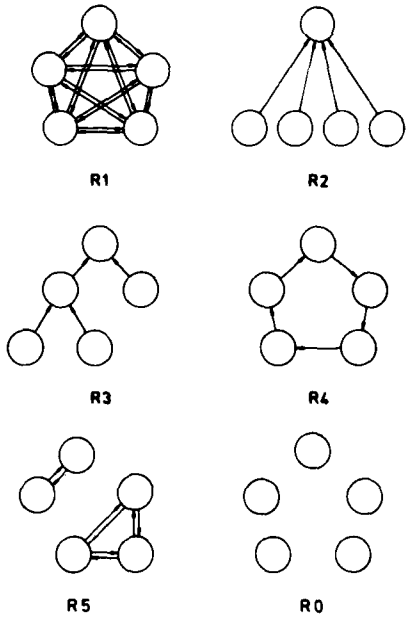


FIG. 15. Types of networks. *Circles*, processes; *arrows*, directed relations.

Modifications in its structure or constituent elements will result in differences in its performance. The performance features reported here reflect the essential properties of its structure, as well as some of its minor details.

Performance

Obviously, the processes of the NORP do not stay internally synchronized under all conditions (Fig. 13, bottom), whereas it is highly likely that under certain conditions internal synchrony is sustained, i.e., when each process is related to every other process (R1) and the variance of the holding time is small. Let us assume that we observe the system for a time interval that is long enough to contain a series of 100 cycles. Because the initial state of the NORP lies always in the predetermined pattern of temporal coordination, a loss of internal synchrony will become apparent only after some time. Thus, we may ask how the length of the series of cycles during which asynchrony does not occur depends on the structure of the network and on the variance of the holding times.

Sustainment of Internal Synchronization

In the case represented in Fig. 16, the NORP structure is R1, and $\sigma_x = 1.6$. The system of 10 processes stays internally synchronized during the entire

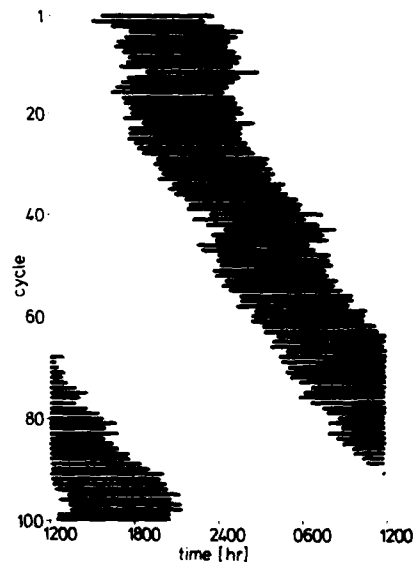


FIG. 16. Chronobiological standard plot of 100 cycles of a NORP equipped with network R1.

series of 100 cycles. Figure 17 shows how the length of the series of synchronized cycles depends on the NORP structure and the standard deviation of the holding time. With networks R1, R2, and R3 (Fig. 15), the system stays internally synchronized for 100 cycles and more, even if the standard deviation of the holding time assumes large values. In case of a cyclic network (R4), there is a certain probability that from time to time the system temporarily loses its coordination. But this is only a transient phenomenon, and the system resynchronizes again after another cycle. A completely different performance is found when there are no interconnections between the processes (R0). In this case, the system gradually runs out of synchrony right from the beginning. However, the desynchronization becomes apparent during a series of 100 cycles only when the standard deviation of the holding time is at least 0.8 hr.

Let us assume that we are investigating NORP like a behavioral scientist who investigates circadian rhythms under constant conditions. We decide to observe the system for a series of 50 cycles. In case the processes are unrelated (R0), the asynchrony will become apparent only if $\sigma_x > 1.4$, i.e., if the participating processes are quite imprecise, with $v_\tau > 0.04$. For smaller σ_x values, however, we might wrongly conclude that the processes are mutually coordinated. This consideration shows that long series of cycles may be observed before a latent asynchrony becomes apparent.

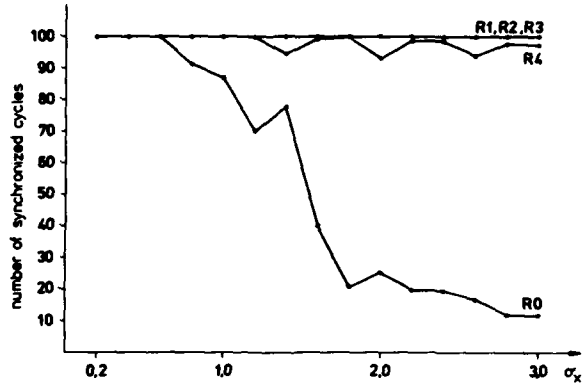


FIG. 17. Persistence of internal synchronization in a NORP: x axis, standard deviation of the holding times σ_x; y axis, length of series of cycles before asynchrony becomes apparent; R0, . . . , R4, types of networks.

Cycle Duration

The results reported thus far have demonstrated a dependence of the performance on the type of the network. Now another aspect of the performance with the networks R1 and R2 (Fig. 15) will be demonstrated (Fig. 18). Let us consider a system containing only one renewal process. Here, the expectation of the cycle duration $\mu_\tau = 100$. The standard deviation σ_τ depends on the standard deviation of the holding time σ_x, as shown in Fig. 17. We can now compare the performance of the NORP containing 10 processes with this case. The standard deviations s_τ estimated for a NORP with a structure R1 or R2 are drastically smaller than σ_τ for a single renewal process. Thus, the interactions of the 10 processes participating in the system cause a strong reduction in the variability of the cycle duration. Moreover, the results show that quite differently structured sets of interconnections can reduce s_τ, i.e., increase the cycle-to cycle precision of the entire system.

There is, however, another aspect of the results shown in Fig. 17. Whereas for R2 the mean value of τ is close to 100 for all values of σ_x, there is a trend of $\overline{\tau}$ for R1. This effect is caused in the following way: Assume that a certain process has just changed its state and is now looking at the processes to which it is related. There are three possible cases: (a) The looking process realizes that it is on schedule. In this case no action is initiated. (b) It notices that it is ahead of schedule. In this case, it will slow down a little by augmenting the expectation of the holding time of the just-assumed state. (c) It notices that it is behind schedule.

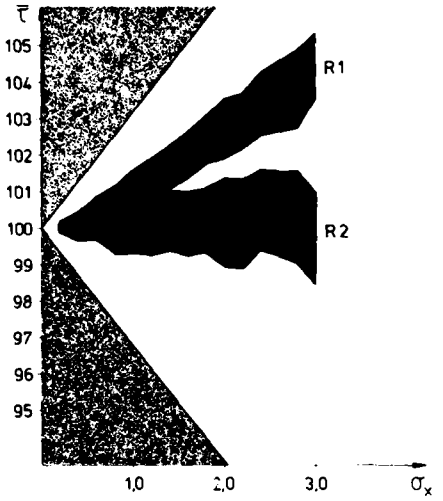


FIG. 18. Mean period $\bar{\tau}$ and standard deviation of the period s_τ of a NORP in internal synchronization: x axis, standard deviation of the holding times; y axis, mean period. *Lighter shading*, region $y:\{y > \mu_\tau + \sigma_\tau\} \cup \{y < \mu_\tau - \sigma_\tau\}$ for a single renewal process; *darker shading*, $y:\{\bar{\tau} + s_\tau > y > \bar{\tau} - s_\tau\}$ for NORPs with networks R1 and R2, respectively.

Of course, it will speed up now. But at the same time there is another effect: During this delay, some other processes may have looked at the process and concluded that they were relatively ahead of their schedules. Consequently, they have slowed down. Thus, there are several slowing-down actions and just one speeding-up action. The two antagonistic types of actions are not balanced, and the probability for an increase of the overall mean cycle duration is greater than the probability for a reduction in this case. This shows that a NORP with network R1 will increase its mean period with increasing probability of control actions, i.e., increasing variance σ_x^2 of the holding time.

Imprecision of the Cycle

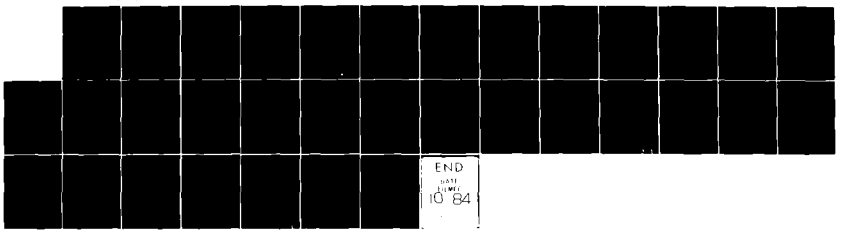
The chronobiological standard representation of a series of simulated cycles of the NORP system (Fig. 16) has a superficial similarity to human rest–activity data. The last aspect of the performance of the model to be discussed here concerns the cycle-to-cycle precision of the simulated data in comparison with the precision of the human circadian clock.

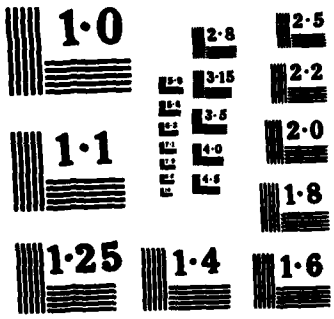
An analysis of simulated data obtained with the networks R1 and R2 by the procedure described earlier yielded the following results: The serial correlation coefficient between consecutive cycle durations is almost always negative. Therefore, the variance of the cycle duration s_t^2 can be split into a component s_κ^2 and a component s_τ^2. The latter is an estimate of the imprecision of the clock-cycle duration. The obtained values of v_τ range from about 0.005 to 0.020 for R1.

These findings lead to the following conclusions: First, series of cycles of

AD A145 712 MATHEMATICAL MODELS OF THE CIRCADIAN SLEEP WAKE CYCLE (U) HARVARD MEDICAL SCHOOL BOSTON MA DEPT OF PHYSIOLOGY AND BIOPHYSICS M C MOORE EDE 01 MAY 84 3/3

UNCLASSIFIED AFOSR TR 84 0781 AFOSR 81 0133 F/G 6/16 NL




1·0
1·1
1·25
1·4
1·6
2·8
3·15
3·5
4·0
4·5
2·5
2·2
2·0
1·8

NORP cannot be described by a renewal model because of the negative correlation of consecutive cycle durations. Second, the cycle-to-cycle precision of the system investigated here is greater than the cycle-to-cycle precision of the human circadian clock estimated from rest–activity data. This may be regarded as a hint that the hypothesis stated earlier, according to which the imprecision of the clock cycle is a result of variability of the periods of the processes constituting the clock, is insufficient to explain the degree of imprecision seen in the human clock cycle in the state of internal synchronization. Thus, influences from shift-inducing factors on the clock must be assumed in order to explain the observed magnitude of imprecision in human observational data.

Although the structure of systems like NORP may be purely speculative with respect to biological reality, the aspects of model performance reported here are intended to emphasize that studies of such systems may help to clarify ideas and support the conception of hypotheses that can lead to new perspectives on circadian phenomena.

CONCLUSIONS

Two properties of the model need to be considered again: NORP does not possess a central unit to which all its elements are connected. The structure of the interconnections is a network solely based on relations between single processes. Thus, NORP does not contain a clock, but as a whole it behaves like a clock. Another main feature of the model is the assumption of memories assigned to each participating process. Here, only the case of predetermined memory contents has been considered, i.e., the possibility of studying learning processes has not been used thus far.

The concept of memory incorporated into the model has still another implication: A process that is just looking at another process has to collect and interpret information about its own state and the state of the process it is looking at. This information then is compared with the respective piece of information in its memory. As a result of the comparison, some control mechanism is activated, if necessary. This mechanism may, in principle, work on a subconscious level, or, in humans, it may be a conscious cognitive process. In the latter case, some specifically human abilities may influence the mode of operation of the mechanism. Two factors may be particularly important: They are goal-oriented activations of control mechanisms (C. Eastman, *this volume*) and misinterpretations of information about the present state of certain body functions. These kinds of factors have not yet been explored systematically by incorporating them into models of the circadian system, either in the empirical domain or in the theoretical domain.

Different kinds of biological systems that possess the ability to generate circadian rhythms are known. The existence of cyclic biochemical processes with circadian period in single cells has been demonstrated, and these are self-sustaining under constant conditions. Also, the ability to generate persisting

circadian rhythms has been proved for delimited structures of the central nervous system in higher vertebrates. Besides these two layers in the hierarchy of life functions, other systems of functions may contribute to sustainment of circadian rhythms under constant conditions. In complex organisms, especially in humans, these different systems supporting the persistence of circadian rhythms operate simultaneously in a hierarchical structure. Cyclic processes in single cells constitute the elementary pacers in a neuronal network. The network as a whole then exhibits circadian rhythms. These signals control various physiological functions such as the production rate of hormones and the body temperature. The behavior of animals or humans is influenced by these physiological processes, and vice versa.

The processes in single cells, as well as on the level of neuronal networks, are being explored experimentally and have been described formally by appropriate mathematical models. The possible contribution of the layer of physiological and behavioral variables to the sustainment of circadian rhythms is still not well understood. The purpose of the model suggested here is to provide a means for exploration of the performance of a complex system of simultaneously operating interconnected processes that shares some basic structural features with the layer of physiological and behavioral variables. Cognitive processes, especially storage and retrieval of information in memory, constitute essential components of the model.

A main result obtained from the Monte Carlo studies is the fact that already the availability of limited information about the familiar pattern of temporal coordination of life functions from a memory constitutes a factor that supports the sustainment of mutually synchronized circadian rhythms under constant conditions. This fact appears to be especially relevant for an understanding of the phenomena occurring in isolation experiments with humans.

REFERENCES

1. Borbély, A. A. (1982): A two-process model of sleep regulation. *Hum. Neurobiol.*, 1:195–204.
2. Czeisler, C. A. (1978): Human circadian physiology: Internal organization of temperature, sleep-wake and neuroendocrine rhythms monitored in an environment free of time cues. Dissertation, Stanford University.
3. Dirlich, G., Kammerloher, A., Schulz, H., Lund, R., Doerr, P., and von Zerssen, D. (1981): Temporal coordination of rest-activity cycle, body temperature, urinary free cortisol and mood in a patient with 48-hour-unipolar depressive cycles in clinical and time-cue-free environments. *Biol. Psychiatry*, 16:163–179.
4. Enright, J. T. (1980): *The Timing of Sleep and Wakefulness*. Springer-Verlag, New York.
5. Enright, J. T. (1981): Methodology. In: *Handbook of Behavioral Neurobiology*, Vol. 4, edited by J. Aschoff, pp. 11–39. Plenum, New York.
6. Feller, W. (1958): *An Introduction to Probability Theory and Its Applications*. Wiley, New York.
7. Lehmann, U. (1976): Stochastic principles in the temporal control of activity behavior. *Int. J. Chronobiol.*, 4:223–266.

8. Lehmann, U., Neumann, D., and Kaiser, H. (1974): Gezeitenrhythmische und spontane Aktivitaetsmuster von Winkerkrabben. I. Ein neuer Ansatz zur quantitativen Analyse von Lokomotionsrhythmen. *J. Comp. Physiol.*, 91:187–221.
9. Lund, R. (1974): Circadiane Periodik physiologischer und psychologischer Variablen bei sieben blinden Versuchspersonen mit und ohne Zeitgeber. Doctoral dissertation, München.
10. Pavlidis, T. (1973): *Biological Oscillators: Their Mathematical Analysis*. Academic, New York.
11. Pittendrigh, C. S. (1965): On the mechanism of the entrainment of a circadian rhythm by light cycles. In: *Circadian Clocks*, edited by J. Aschoff, pp. 277–297. North Holland, Amsterdam.
12. Pittendrigh, C. S., and Daan, S. (1976): A functional analysis of circadian pacemakers in nocturnal rodents, I. *J. Comp. Physiol.*, 106:223–252.
13. Wever, R. A. (1979): *The Circadian System of Man*. Springer-Verlag, New York.
14. Winfree, A. T. (1980): *The Geometry of Biological Time*. Springer-Verlag, New York.
15. Zulley, J. (1979): *Der Einfluss von Zeitgebern auf den Schlaf des Menschen*. R. G. Fischer Verlag, Frankfurt.

DISCUSSION

Dr. Kronauer: In the analysis of the synchronized human data, did you assume that they were random according to the renewal process when you analyzed them? Or did you conduct some test of this?

Dr. Dirlich: I did not assume that they were random. The procedure that I used does not discriminate between systematic and random variations of tau. It just is a procedure that separates the variance component caused by the coupling and the variance component caused by the sloppiness of the clock cycle. But I cannot say if the estimated variance of the clock cycle is due to systematic or random variations.

Dr. Enright: You emphasized in your chapter the importance of models that offer predictions. What kinds of predictions did you arrive at from your model?

Dr. Dirlich: None, at the moment. It is pure speculation.

Dr. Edmunds: I guess there is no point in asking why you think the clocks in unicellular organisms are so precise. The latest *Gonyaulax* work shows a standard deviation in precision of 17 min per day. Would you care to speculate on the difference between sloppy human clocks and the precise unicellular clocks?

Dr. Dirlich: I cannot say much about it. I think there may be a circadian clock in humans of the kind that Dr. Enright has described. If this is the case, it is part of a higher kind of control mechanism. This control mechanism can then modify the speed of the clock at random.

Mr. Pilato: You mentioned that with the hierarchical model, you have got a rather constant tau which was independent of the standard deviations, whereas with the collective model, tau increased with standard deviation. In Dr. Kronauer's two-oscillator model, one has to change tau over time in order to get a progressive change in states, leading to desynchronization. The underlying tau gradually increases, beginning at release from Zeitgeber. At that time, I suspect there may be either a gradual change from hierarchical to collective organization or an increase in standard deviation over time.

Mathematical Models of the Circadian Sleep-Wake Cycle, edited by M. C. Moore-Ede and C. A. Czeisler. Raven Press, New York © 1984.

Exploratory Data Analysis: Published Records of Uncued Human Sleep–Wake Cycles

Arthur T. Winfree

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907; and Institute for Natural Philosophy, West Lafayette, Indiana 47906

Everything should be made as simple as possible . . . but not simpler.
Albert Einstein

I am exploring that limit, in the case of modeling human circadian rhythms, by taking the risk that I may err on the side of oversimplification. I shall attempt to redescribe some recordings of human sleep and waking made in several laboratories in the past decade. My objective is to avoid models as far as possible, then to compare only the sketchiest versions of them against published data in order to find out what further elaborations the data seem to require. A first look at the data suggests that under some conditions there may be an interesting discontinuity in the dependence of wake-up time on prior sleep-onset time. This feature may find simple interpretation in terms of circadian variation in a threshold process that initiates wake-up. It may therefore help to clarify the connection between the continuous variables of oscillator models and the discrete events of sleep onset and wake-up. Also, I have observed none of the regularities in the timing of sleep onset that would be expected according to the dozen or so models currently under consideration.

Curious and elusive regularities lurk in records of sleep and wake transition times. Among the most intriguing are the following:

1. Aschoff and Wever (9, and citations therein) have observed that sleep and wake timing need not stay synchronous with the steadier beating of our circadian clock, as reflected in the ups and downs of core temperature.
2. The more recent discoveries by Czeisler (1,2) have shown that there is a natural periodic time base to which to refer human sleep–wake transitions, that the time of awakening depends mainly on the time of prior falling asleep when both times are referred to that natural period, and that something about human sleep–wake behavior is slowly changing week after week and month after month during temporal isolation.

3. Kronauer (6,7) has demonstrated that a two-oscillator analogy descended from that of Wever, with several critical mutations, does reproduce many of the curious regularities of human sleep–wake timing.

Because I am not in a position to do my own experiments in this area, I have chosen to begin with the published data. In particular, it struck me that one feature of the data—a feature that I have often seen in my own experiments and in those of others on invertebrates—has not been emphasized in manuscripts available to me at this writing. Because it may have an important bearing on contemporary interpretations, I would like to draw attention to it here.

WHAT COMES DOWN NEED NOT HAVE FIRST GONE UP

This feature emerges most starkly in plots of sleep duration against the time of sleep onset, modulo Czeisler's discovered time base. What Czeisler found is that there is a period in the range of 24 to 25 hr, quite sharply defined in some data sets, wherein the sleep onset phase predicts sleep duration with minimal variance. In at least one other data set that I have examined (8, Fig. 8), the same holds true, except that the underlying period must be assumed to increase very slowly during the months of observation. That period turns out to be the dominant component of core-temperature fluctuations, too; so a plot of sleep duration against the phase of sleep onset in the temperature rhythm also tends to look quite orderly. So does a plot of duration against sleep onset time measured in hours past the most recent temperature minimum or since the middle of the sub-average-temperature interval, etc. Figure 1 shows a series of plots derived from the data of Jouvet et al. (4) in the manner of Czeisler et al. (2): each sleep duration is plotted against its time of onset modulo and assumed period, and then the period chosen is varied by 2 min from one plot to the next until a period is found (a fixed period in this case) that brings out a functional dependence. In this case, the data look most orderly at $T = 24$ hr, 20 min.

The most conspicuous feature of this plot is its lack of continuity; it seems to break near a critical phase at which sleep may continue either for an uncommonly short time or for an uncommonly long time, but seldom in between, as observed in less exaggerated form by Czeisler et al. (1,2) in PRO 1 and other subjects. Should one draw through these data an idealized smooth curve? To me, these data and those of Zulley (13), plotted by reference to the daily temperature minimum, suggest an alternative piecewise-continuous redescription in the fashion of Fig. 2.

It is not necessary in theory that the curve through such data should rise and fall continuously, as, for example, in the work of Kronauer et al. (6, Fig. 10). A discontinuity (even an overlapping discontinuity, as fantasized in Fig. 2) can be a sensible alternative interpretation of the observed sleep and wake times, as I shall show below.

The simplest model provides an acceptable interpretation of discontinuity in

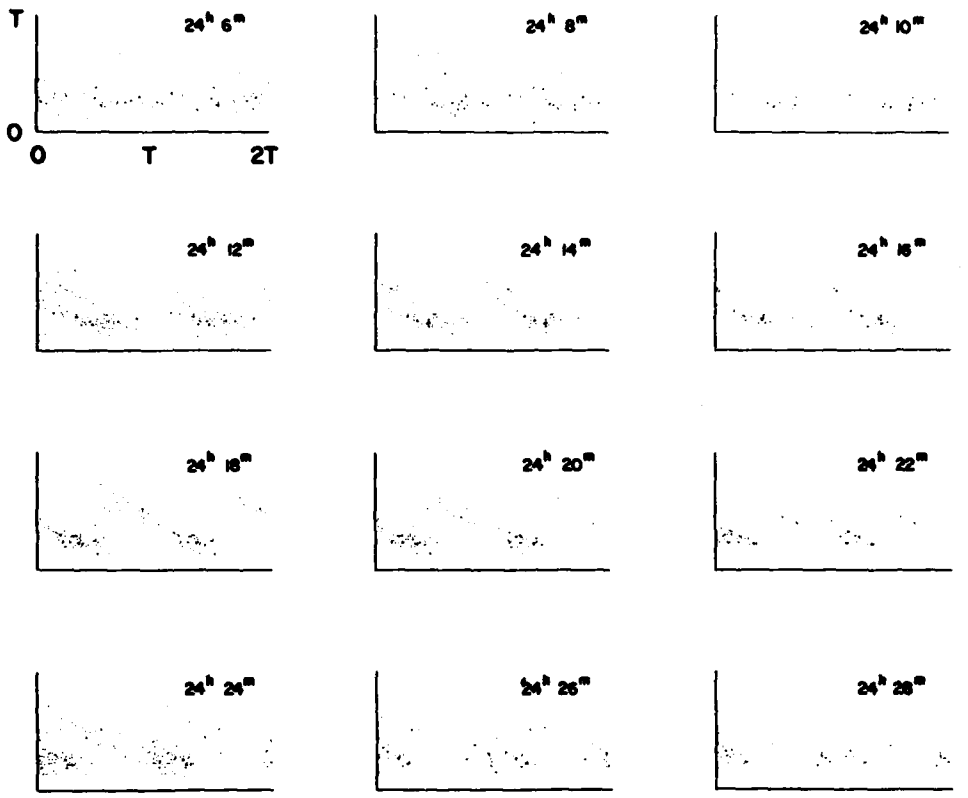


FIG. 1. Sleep duration is plotted against arbitrary phase in a cycle of duration T throughout the free-running segment for subject JC (4, Fig. 2). Twelve plots are shown, at T values increasing by 2 min from 24 hr, 6 min to 24 hr, 28 min. Czeisler optimized T (1, p. 289) by minimizing the variance of durations averaged over onset time modulo T. His procedure computes a large variance at any T that shows parts of the data cloud overlapping at a discontinuity. Suspecting that such might actually occur, we chose instead to "eyeball" the plots for best confinement of data to a one-dimensional locus in the plane. This alternative procedure has the advantage of detecting regularity without *a priori* excluding T values that reveal bimodal distribution of sleep durations.

sleep durations. Suppose for the moment that wake-up is initiated when core temperature rises to a certain point (call it phase 25 on a scale 0–25) in its daily cycle. Then sleep duration is simply 25 minus the time of sleep onset in this cycle. At phase 25, duration jumps discontinuously from 0 to 25. This model is, of course, too simpleminded. But now suppose I tend to wake earlier than the standard phase if already well rested, and to linger longer asleep if I have only recently retired when phase 25 comes due. Then the discontinuity is reduced toward the 8 to 12 hr more typically observed. Biological variability might make durations bimodally unpredictable near the critical phase, thus providing the appearance of overlap. This kind of model runs afoul of data in other essential

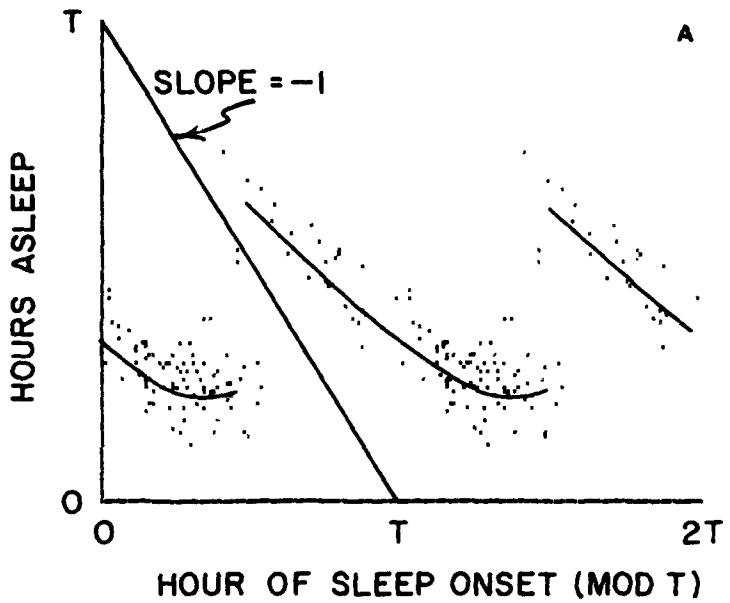


FIG. 2A. One panel of Fig. 1 is overlaid with a suggested curve threading the data. The curve is discontinuous and lacks upslope in the manner described earlier (10).

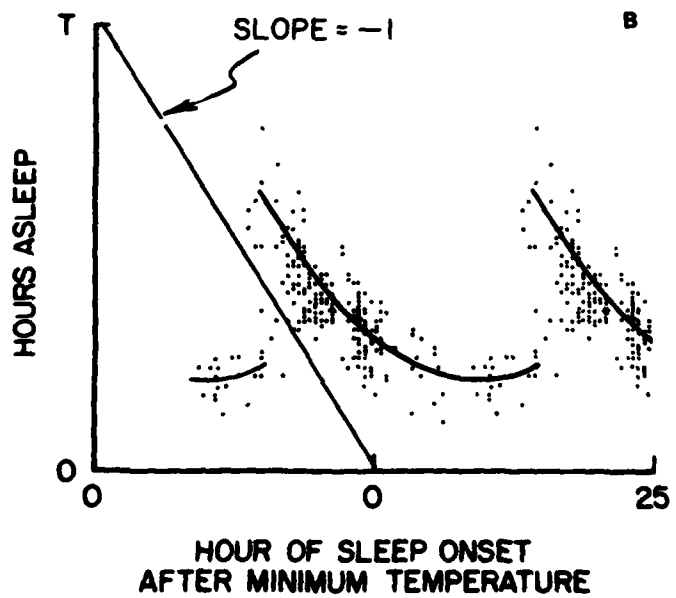


FIG. 2B. The data of Zulley (13) are similarly overlaid. These points were located horizontally by absolute hours after the most recent temperature minimum, not by simply postulating a rhythmical influence and then seeking its period as in Fig. 1. These two methods amount to nearly the same thing if the postulated influence is conspicuous in temperature fluctuations. However, Zulley's data are probably somewhat displaced by the temperature-reducing effect of sleep. The diagonal line (hour of onset + hours asleep = 0) represents wake-up at the temperature minimum. Very few wake-ups occur in this region.

regards, but it suffices here to make the point that there is nothing intrinsically unreasonable about discontinuity in the dependence of sleep duration on sleep onset phase.

Czeisler (1) also noted the sudden change in sleep durations and compared it to the "breakpoint" in the transients of phase resetting. Kawato et al. (5) presented an interpretation according to which an underlying mechanism phase-shifts abruptly to change the sleep duration. My own preference is for an interpretation, sketched below, more in line with the models of Wever (9; and *this volume*) and Kronauer et al. (6; and *this volume*). But the point I wish to make here is that there is no *theoretical* reason prejudicial to interpretation of the data by a piecewise-discontinuous curve or (what is the same) recognition of a range of circadian clock phases during which wake-up seldom or never occurs. This gap is conspicuous in the raster-plotted raw data (when the raster period is accurately matched to the individual's circadian clock) as a vertical band devoid of spontaneous awakenings (11,12).

I have been cautioned (R. Kronauer, *personal communication*) that although discontinuities may present no obstacle to model makers, they may still be artifacts of data processing:

1. There are *some* dots that might indicate a very steep upslope rather than strict nonexistence of any curve rising through intermediate sleep durations.

2. In sleep–wake recordings unaccompanied by temperature data, and to a lesser extent even with such data, there is some slight latitude of choice about the base period. It commonly happens in long records that the longest sleeps occur preferentially near the end of the record. By choosing a slightly longer base period, one can plot these sleeps at earlier phases, overlapping short sleeps initiated at the same phase earlier in the record. This effect spuriously steepens the upslope or introduces frank overlap of longest and shortest sleeps.

3. A different problem afflicts analysis of data in which base period seems unsteady. In such cases one can only plot sleep onsets relative to the recent temperature excursions. Wake-up *causes* a temperature rise; so the temperature minimum cannot easily *follow* wake-up. Thus emerges an apparent but artifactual gap in wake-up phases, and the corresponding discontinuity in sleep durations. This effect would tend to condense sleep onsets around one quarter cycle before apparent temperature minimum while rarefying wake-ups in that range, with the opposite effect near the opposite phase in both cases. This is exactly what Zulley's data show (Fig. 2B). This "masking" of the core-temperature rhythm by the direct effect of activity can be compensated numerically, but I have not yet seen sleep–wake timing replotted on this basis. In principle, this is a continuous distortion; so it cannot make a continuous curve look discontinuous, or vice versa. But, in fact, one deals with data points, and only *imagines* curves. I am not sure how one's imagining might be affected after making this correction.

4. It would appear from the work of Czeisler et al. (2) and Kronauer et al. (6)

that bimodality and discontinuity are not universal: A smooth, almost sinusoidal, curve derived from the two-oscillator model fits quite nicely to the data of PRO1 (6, Fig. 10). The models computed by Wever and by Kronauer (at least with parameters as then adjusted) do not exhibit discontinuity. As illustrated by Fig. 10 of Kronauer et al. (6), every realizable sleep duration can be elicited at either of two phases of sleep onset, one where the curve is *descending* and one in the midst of its opposite ascent.

Those are some of the caveats. Only laboratory observation will clarify the matter; so it may be years before it is known how common this feature is that can now be perceived only ambiguously. But suppose for the moment that this discontinuous interpretation may be as appropriate as the original smooth interpretation and enquire what it may signify. The question is of interest in itself, concerning, as it does, a distinct qualitative feature of the psychophysiological mechanisms of sleep. It also has interest in connection with models. According to one mathematical metaphor, any mechanism involving accumulation or depletion of some quantity toward a threshold that is subject to a periodic influence can exhibit such bimodality, overlap, and discontinuity. I shall present a simple caricature to illustrate how this comes about and to underscore some of the diagnostic idiosyncrasies of this class of mechanism.[1]

I believe attention to this feature might prove useful to discriminate between admissible and inadmissible ways of relating the smoothly varying quantities of an underlying oscillator to the discrete transitions between sleep and waking. In underscoring the deficiencies of Wever's early choice of a threshold interpretation, Kronauer et al. (6) pointed to the need for such discriminating observations.[2]

A GATING MODEL

Kronauer et al. (6) and Aschoff and Wever (9, and citations therein) argued that human sleep–wake data are best construed as revealing two separate circadian clocks: one of relatively constant 24- to 25-hr period and another of longer and more labile period. [Note, however, the dissenting opinions of Dirlich (*this volume*) and Eastman (*this volume*).] They further inferred that both oscillators involve two or more smoothly varying quantities in their basic mechanisms and that both oscillators have limit cycles to which they recover at a leisurely pace during transient changes of amplitude. This understanding of

[1] Daan and Beersma (*this volume*) have described a well-developed model of just this sort. It goes further to also interpret wake durations in the same terms, despite the lack of corresponding regularity in the published observations (*vide infra*).

[2] I am indebted to R. Kronauer for his subsequent observation that with appropriate adjustment of its threshold mechanism, his model can also represent the discontinuity, as shown in Kronauer et al. (7, Fig. 10). Figure 3 from the same work reinterprets the PRO1 data as in my Fig. 2 and the adjusted model.

human clockworks needs to be set in context by understanding also why nearby alternative interpretations do not work. As a step toward verifying both the sufficiency and the necessity of the existing two-oscillator model, I would therefore like to gather observations appropriate for testing it against the many alternatives it seems to be provoking.

I now present one more such alternative: a simple interpretation of the dependence of waking time on prior sleep onset, both being measured by reference to a smoothly varying rhythmic influence. This influence may be construed as coming from the autonomous circadian clock that also influences core temperature or, when external scheduling is dominant, as coming from some outside Zeitgeber. Suppose that the first oscillator, the one with steadier period, which is more conspicuous in core temperature, makes its appearance in the sleep–wake system simply as a smooth rhythmical modulation of some threshold at which wake-up is initiated (Fig. 3). (There is no cause to assume that its maximum and minimum correspond to those of core temperature.) This interpretation is lifted directly from my interpretation of the pupal eclosion rhythm in the fruitfly (10, pp. 403–406). The model has been studied in some detail mathematically, but, skipping all that, the bottom line is that models of this class exhibit a characteristic idiosyncrasy that seems to appear also in the dependence of sleep duration on sleep onset time (Fig. 1).

Figure 3 illustrates the timing of wake-up given one particular choice of sleep onset time, where the rising quantity (call it "restedness") starts its rise from

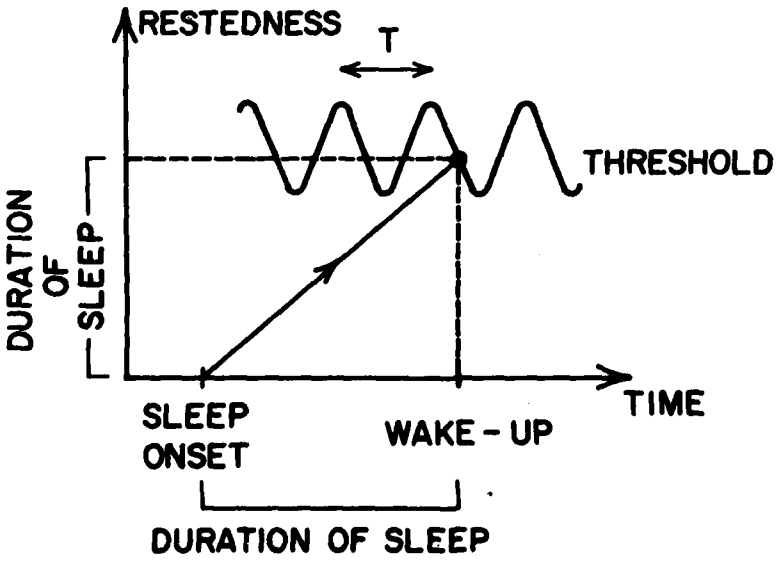


FIG. 3. "Restedness," defined proportional to cumulative sleep, rises toward a threshold at which wake-up is irrevocably initiated. This threshold is postulated to vary with subjective circadian time at a period *T* to be discovered in the data.

baseline. (It would, of course, be equivalent to fantasize a declining "tiredness.") Were there no modulation of the threshold level, sleep duration would be the same no matter when sleep begins; in fact, there would be no grounds for distinguishing one "when" from another, so far as the threshold is concerned. With nonzero threshold modulation, the duration of sleep depends on the timing of sleep onset. In fact, with arbitrarily strong modulation, the dependence is exactly as in the "simplest" model described earlier: Wake-up can occur only at a standard phase in this rhythm. So one needs to consider modulation at some moderate amplitude between the extremes.

In Fig. 3, sleep duration would decrease if the phase of sleep onset were moved to the right. Moved farther right to a critical phase, the rising quantity would altogether miss the trough of the threshold rhythm, and hit threshold only much later (Fig. 4). (The same might be accomplished by sleep deprivation, causing restedness to start from a lower origin or, equivalently but more conveniently for graphic anticipation of the results, elevating the threshold curve.) At this critical phase, the shortest sleep abruptly changes to the longest. *There is no ascending branch of the curve.* In fact, if very short sleep sometimes fails to terminate at threshold, i.e., if sleep sometimes lingers until the threshold is uncrossed again a short time later (Fig. 5), then shortest and longest may be two alternative choices for the duration of sleeps initiated near that critical phase. The top and bottom parts of the curve would then overlap in the way suggested by the data of Fig. 1. Wake-up is gated by the rise and fall of the threshold. It can occur spontaneously only in a certain range of phases (where the threshold is not rising too fast). Forced waking at other times might be uncomfortable and unproductive.

At sufficiently high amplitudes of threshold modulation, wake-up never occurs in a certain range of phases in the temperature cycle, regardless of onset phase. But at sufficiently *low* amplitude of modulation, wake-up can occur at *any* phase, depending on onset phase. In other words, if the threshold modulation is weak (if the 25-hr circadian rhythm and/or exogenous Zeitgeber has less impact on the sleep–wake mechanism), then no discontinuity occurs. This is because the weakly modulated threshold never rises faster than restedness (Fig. 6). In such a case, duration varies smoothly with onset phase, as represented by Kronauer et al. (6, Fig. 10). In this situation the duration of sleep averages much longer, too, corresponding more nearly to the time it takes for restedness to reach average threshold level, never being interrupted by a plunging descent of threshold. Could this have anything to do with Wever's observation (9) that when a subject's temperature amplitude decreases, his sleeps average longer?

This caricature of the rhythmically modulated timing of a discrete event can be put together in a single three-dimensional diagram (10), as in Fig. 7. This particular diagram is computed from a mathematically tidy representative case in which restedness increases steadily (by definition) while the threshold for wake-up *fluctuates sinusoidally* (or in any other fashion). Note that the

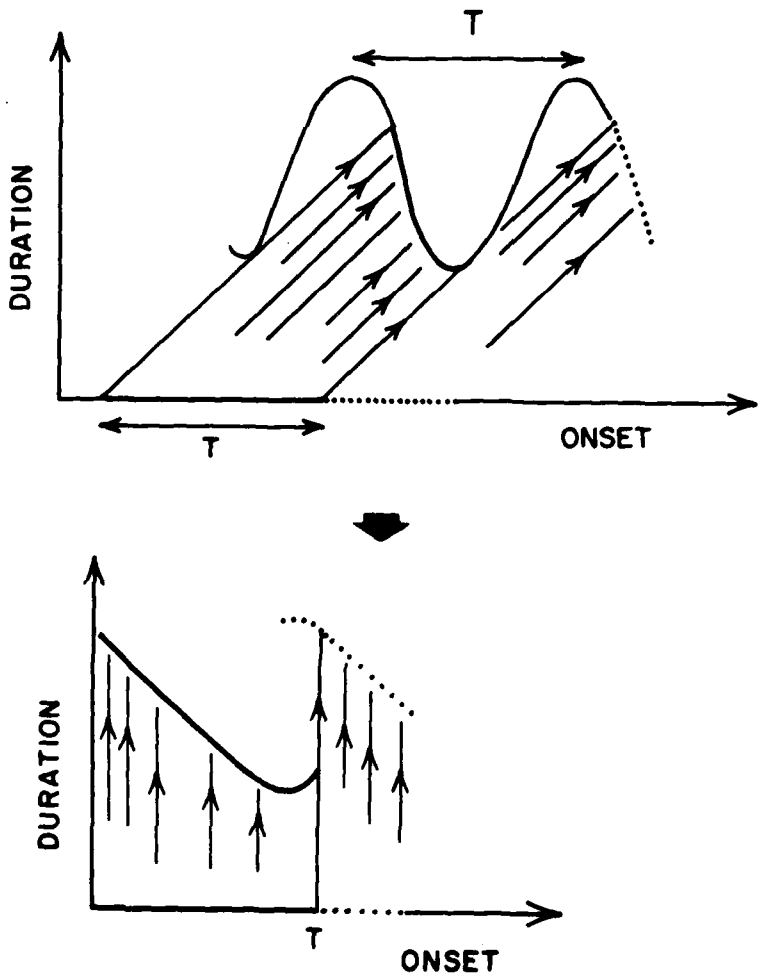


FIG. 4. Top: Figure 3 is filled in with threshold interceptions from all sleep onset phases spanning 1.5 cycles of *T*. **Bottom:** The dependence of duration on choice of onset is replotted in upright coordinates.

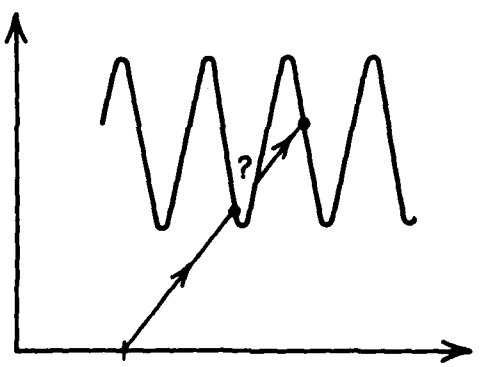


FIG. 5. As in Fig. 3, but differently scaled to suggest how a bimodal distribution of sleep durations might arise by transgressing the threshold only briefly near a critical phase.

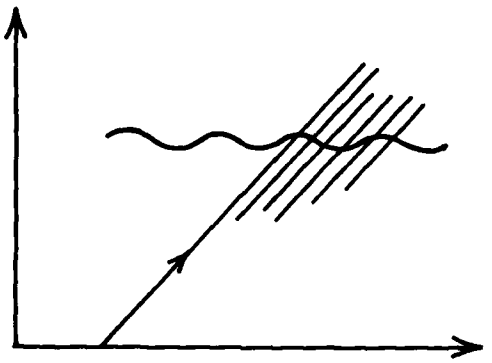


FIG. 6. As in Fig. 3, but differently scaled to suggest how weak modulation of threshold may result in a smooth up-and-down dependence of duration on onset phase without a break, as in PRO1 (1,2,6).

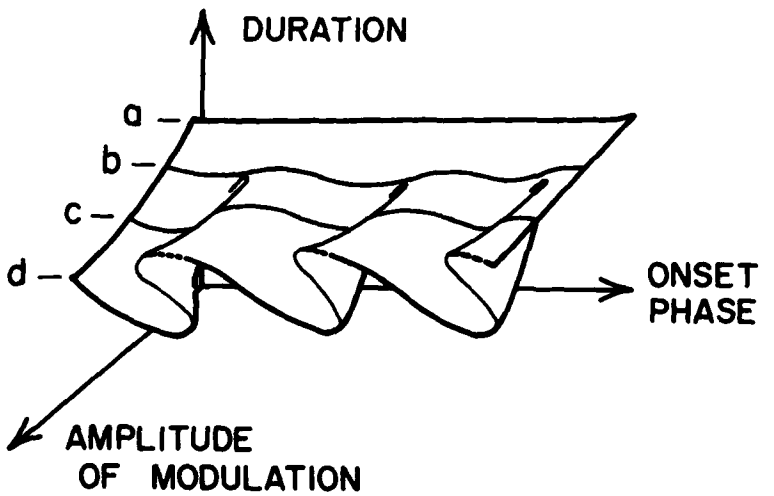


FIG. 7. Threshold-modulation models produce different curves of sleep duration versus sleep onset according to the amplitude of modulation (by an endogenous clock or by an external scheduler). At amplitude 0, duration is always the same (*a*). At low amplitudes (*b*) it varies somewhat according to onset time modulo *T*. At middling amplitudes (*c*) it gets a steep narrow upslope region of high variance, but durations are still unimodally distributed. At high amplitudes (*d*) the upslope has vanished, leaving a frank discontinuity, even with overlapping edges. Durations are bimodally distributed near this phase. The whole picture is called a "Whitney cusp." It is typical of a great variety of control systems.

threshold could equivalently be regarded as constant while restedness rises sometimes faster and sometimes falls back a little; the consequences for timing would be identical. In fact, any model exhibits this qualitative behavior if it combines a smoothly rhythmical wiggle with a quantity slowly approaching a threshold to switch sleep on or off. In particular, the oscillator models of Wever (9) and Kronauer (6) should show such duration-versus-onset curves when parameters are so adjusted that the wiggle's amplitude exceeds the cusp point in Fig. 7. Unfortunately, this statement seems to apply equally to sleep duration (wake onset) and wake duration (sleep onset) in those models.

Such notions can be put to trial by sleep-deprivation experiments and/or by changing the vigor or period of rhythmic driving (the threshold modulation). The results expected (discontinuities appearing or disappearing, moving right or left) would be hard to mistake in the laboratory.

Before going on, note that if a process similar to that here conceived to govern sleep duration were at work also in the determination of wake duration, then my simplified caricature would constitute a second oscillator functionally equivalent to the y (activity) oscillator of Kronauer et al. (6). [This is the model of Daan and Beersma (*this volume*).] It differs from his only in that my amplitude is inflexibly determined and my equilibrium state either does not exist or is a violent repeller. This is getting to be fairly exotic speculation, but the business can, in principle, be resolved by a singularity experiment (10). This is not yet technically feasible in humans, so there is not much point in belaboring it. But when and if one does find out how to perturb human clocks to their phaseless states, then Kronauer's model can be expected to show that recovery will be rather slow; if so, then simpler models of the sort here explored will be starkly inadequate.

For the present, I have not been able to make sense of wake-duration data (i.e., the times of sleep onset), so I postulate no second oscillator. This chapter is confined to an interpretation of sleep duration alone, or, in other words, of wake-up timing.

WHAT DETERMINES SLEEP ONSET?

Sleep durations taken from any stage in the secular progression of sleep–wake behavior under temporal isolation all seem to fall on the same curve, as though the shape of Fig. 7 and the modulation amplitude both remain unchanged. While the patterning of sleep–wake alternations is metamorphosing, the timing of sleep is not. The change, then, may lie exclusively in the mechanism determining sleep onset, or the duration of waking. I need now to direct my exploratory data analysis to the other half of the sleep–wake cycle by asking "What determines when one goes to sleep?" Is the same periodic time base a conspicuous organizer of that dependence? Does the time of sleep onset depend

mainly on the previous time of waking or mainly on the previous times of waking and sleep onset (perhaps through their difference, the previous sleep duration), or additionally on more remote prior events?

Perhaps it would not be surprising to find disappointingly little determinism here, as we all have the subjective impression that we can voluntarily defer sleep to a much greater extent than we can adjust the moment of waking, given the time of sleep onset. Attempts in this direction may be premature. Certainly they are hobbled still by a shortage of long unperturbed free runs in the published record, and by the intriguing fact that during such runs, some essential parameter is apparently drifting and, as it slowly drifts, distorting the functional dependences I seek to observe. But it would be so valuable to know how one sleep episode (or several) predetermines the onset of the next (if it does) that I have undertaken to spend some time looking. I am inclined to take a lesson from the recent mathematical literature of iterated mappings: that lovely regularities may lurk in ostensibly random data, awaiting discovery by someone in possession of a long enough and steady enough time series who makes a felicitous choice of what to plot against what.

It seems that this can be done in a model-free and theory-independent way. If I had two records of alternating sleep and wakefulness, identically timed through a dozen episodes, then I might reasonably expect the next sleep onsets to be identically timed. Just as a matter of phenomenological description, the hour of sleep onset should be somehow implicit in the timing of earlier events. This is known to be so in the case of wake onset; in fact, only one previous event suffices for excellent prediction. But in the case of sleep onset, I have not found any clear regularity either in terms of the phases of the previous three transitions or in terms of the durations of prior waking and sleep.

In contrast, every model I have examined (by making artificial data sets to submit to the same data-analysis programs) does show clear regularities of diverse sorts, but none of these appear in the data. The regularities in the models are, of course, quite similar for the two halves of the sleep–wake cycle. This is because existing models treat the two stages evenhandedly, as two sides of a coin, e.g., above and below a threshold. So my impression is that timing of sleep and of wake is not so symmetrically handled in real people; for example, whatever is changing during months of isolation appears to affect the timing of sleep onset far more than it affects the timing of wake onset.

ACKNOWLEDGMENTS

All the participants in the symposium that was the basis of this book have been more then generous in bringing me up to date since I first undertook to learn about circadian clocks. Specifically, S. Samn at the USAF School of Aerospace Medicine drew my attention to the practical importance of jet lag. From J. Enright I learned of Wever's book *The Circadian System of Man*, and F. Sulzman directed me to Czeisler's thesis on human circadian physiology. M.

Moore-Ede introduced me to the coupled-oscillator model of Kronauer. Funding for many travels during sabbatical leave from Purdue University came through the Institute for Natural Philosophy and USAF School of Aerospace Medicine.

REFERENCES

1. Czeisler, C. A. (1978): Human circadian physiology: Internal organization of temperature, sleep-wake, and neuroendocrine rhythms monitored in an environment free of time cues. Dissertation, Stanford University Medical School.
2. Czeisler, C. A., Weitzman, E. D., Moore-Ede., M. C., Zimmerman, J. C., and Knauer, R. S. (1980): Human sleep: Its duration and organization depend on its circadian phase. *Science*, 210:1264–1267.
3. Enright, J. T. (1980): *The Timing of Sleep and Wakefulness*. Springer-Verlag, Berlin.
4. Jouvet, M., Mouret, J., Chouvet, G., and Siffre, M. (1974): Toward a 48-hour day: Experimental bicircadian rhythm in man. In: *The Neurosciences: Third Study Program*, edited by F. Schmidt and F. Worden, pp. 491–497. M.I.T. Press, Cambridge, Mass.
5. Kawato, M., Fujita, K., Suzuki, R., and Winfree, A. T. (1982): A three-oscillator model of the human circadian pacemaker controlling body temperature and the sleep-wake cycle. *J. Theor. Biol.*, 98:369–392.
6. Kronauer, R. E., Czeisler, C. A., Pilato, S., Moore-Ede, M. C., and Weitzman, E. D. (1982): Mathematical model of the human circadian system with two interacting oscillators. *Am. J. Physiol.*, 242:R3–R17.
7. Kronauer, R. E., Czeisler, C. A., Pilato, S. F., Moore-Ede, M. C., and Weitzman, E. D. (1982): Mathematical model of the human circadian system with two interacting oscillators. *Am. J. Physiol.*, 242:R3–R17.
8. Mills, J. N., Minors, D. S., and Waterhouse, J. M. (1974): The circadian rhythms of human subjects without timepieces or indication of the alternation of day and night. *J. Physiol. (Lond.)*, 240:567–594.
9. Wever, R. (1979): *The Circadian System of Man*. Springer-Verlag, Berlin.
10. Winfree, A. T. (1980): *The Geometry of Biological Time*. Springer-Verlag, Berlin.
11. Winfree, A. T. (1982): Human body clocks and the timing of sleep. *Nature*, 297:23–27.
12. Winfree, A. T. (1982): The timing of sleep and wakefulness in men and women. *Am. J. Physiol.*, 243:R193–R204.
13. Zulley, J. (1980): Discussion. *Sleep*, 2:344–346.

DISCUSSION

Dr. Wever: When we look for precision in various phase points, sleep onset is by far the least precise reference phase that exists in the sleep–wake cycle; it has 50% larger variability than all other phase point reference points, including, for instance, the midpoint of sleep.

Dr. Czeisler: Are your data that you have presented all from internally synchronized subjects?

Dr. Wever: Yes, only internally synchronized people.

Dr. Weitzman: If you look at what sleep stage ends the sleep episode during free run, as compared to entrained subjects, it is very clear that there is a much increased probability (i.e., about 80%) that the subject will wake up out of REM sleep. However, in the entrained condition, the probability is down to 20 to 30%. That is in spite of the fact that REM sleep during free running is shifting earlier; so there should even be a

lower probability of waking out of that stage. Now, that means that there is some powerful set of presumably ultradian REM/non-REM cycles which is somehow structuring that endpoint. This suggests there is some kind of interrelationship between the ultradian REM/non-REM cycle and the circadian sleep–wake cycle which is very predictable.

Dr. Winfree: That sounds like a factor for determining the endpoint of the REM/non-REM cycle, but I do not believe it to be related to the timing of circadian cycles.

Dr. Rosenthal: I am fascinated by your model.

Dr. Winfree: Don't call it a model! [*Laughter*]

Dr. Rosenthal: I am very interested in how it is similar to Dr. Borbély's model in that it presupposes an interaction between a single circadian oscillating system and some "buildup" or "hourglass" system. I am particularly interested in the way in which it may explain abnormal sleep phenomena such as we see in depression. One could postulate a reduced pressure to sleep, which could explain the reduced delta sleep seen in depression. Your suggestion of a reduced threshold for wakefulness would explain a lot of things.

Dr. Winfree: Well, that is precisely the point.

Dr. Rosenthal: Depressives do have a different sleep duration. They have early morning waking, which could be predicted by your model, but it would be predicted that they would have early morning waking that would cause sleep deprivation. There would, therefore, be REM pressure which might explain why there is the observed increase in REM in the early part of the night when the depressive subsequently goes to sleep. The restriction of REM would increase REM density in the early part of the night. All of these phenomena could be explained in terms of the picture you have drawn.

Mathematical Models of the Circadian Sleep-Wake Cycle, edited by M. C. Moore-Ede and C. A. Czeisler. Raven Press, New York © 1984.

Sleep Duration for Human Subjects During Internal Desynchronization

J. T. Enright

Scripps Institution of Oceanography, University of California, La Jolla, California 92093

Zulley [in Czeisler et al. (2)] and Czeisler et al. (1), in their analyses of data from circadian studies of isolated human subjects, have noted that when a subject shows dissociation of the wake–sleep rhythm from the core-temperature rhythm, there is a systematic relationship between the phase of the temperature cycle at which the subject falls asleep and the duration of the ensuing sleep. On the basis of these data, as well as of his own analyses of similar data, Winfree (*this volume*) has suggested that this quantitative relationship may well involve a discontinuity: If sleep onset occurs at a certain critical phase of the underlying endogenous pacemaker rhythm (as indexed by the temperature cycle), the result can be either a very short sleep or a very long one, with intermediate values being unlikely. A slightly earlier sleep onset leads consistently only to very short sleep, and a slightly later sleep onset only to very long sleep. Both Winfree (*this volume*) and Daan and Beersma (*this volume*) have proposed models, of differing complexity, in which such a discontinuity is predicted to occur because of interactions between a cumulative renewal process and a rhythmically fluctuating threshold. In this chapter I shall outline a possible alternative explanation for that phenomenon.

AN INTERESTING RESEMBLANCE

I have been struck by the similarity between these empirical results and the outcome of certain computer simulations based on models I have called "coupled stochastic systems" (3). Those models were initially formulated as a means of accounting for wake–sleep data from nocturnal rodents and from diurnal birds. The parameter values for the specific model of interest here were assigned to provide an adequate fit to phase-shift data for 6-hr light stimulation of the house sparrow (3, Figs. 11.5 and 11.6). The simulations, in which there are resemblances with the human data, represent predictions for the bird's rhythm when single light stimuli of varying duration are administered during free run, with onset of the light always corresponding to onset of activity (wake-up time) (3, Fig. 13.8). Because of the formulation of this specific model

(corresponding to the behavior of birds), onset of sleep would be expected to occur at the end of the light treatment for all stimuli longer than about 4 hr. When those predictions are replotted in a coordinate system similar to that used by Winfree and others for the human data, the hypothesized behavior of the bird shows a phase-dependent pattern of sleep duration similar to that of the human subjects (Fig. 1). [Note particularly the sudden transition from very short to very long sleeps; compare with human data of Winfree (*this volume*).] Numerous other simulations with this particular model have convinced me that qualitative aspects of this result are due almost exclusively to the time at which the light stimulus ends (3, Fig. 13.6).

DISCUSSION

It may well be that this qualitative resemblance between (hypothetical) bird data and the real human data is only fortuitous; certainly the subsequent behavior of the wake–sleep pacemaker of my models, following such treatment, seems to differ appreciably from that of the body-temperature cycle of human subjects. Nevertheless, it seems worthwhile to pursue briefly the speculation that the resemblance could reflect qualitatively similar processes in the wake–sleep pacemakers of bird and human.

In a bird, the light receptors of importance for circadian rhythmicity are not retinal, but directly in the brain (5); whether the bird is awake or asleep, its

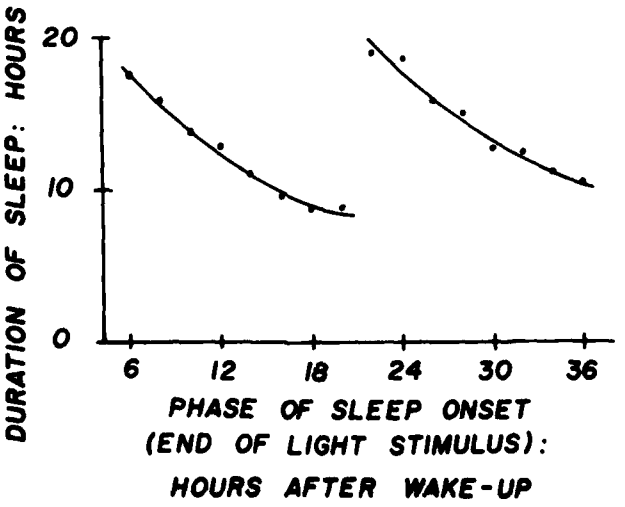


FIG. 1. Predictions for the behavior of a sparrow, following light stimuli that produce onset of sleep at various circadian phases, based on simulations with a coupled stochastic system [replotted from Fig. 13.8 in Enright (3)]. Although the quantitative agreement with the human data (A. T. Winfree, *this volume*) is by no means perfect, it should be noted that none of the values for the model's parameters was assigned with human data in mind.

pacemaker is continually subject to prevailing light intensity (unless, as sometimes happens in dim light, the sleeping bird tucks it head beneath its wing). The pacemaker of a human subject, however, like that of other mammals investigated, can presumably be affected only by retinally perceived light, with the result that when awake, with open eyes, the human administers to his pacemaker system a light stimulus. Onset of sleep and/or turning off of the room lights are the usual ways in which that stimulus is terminated each day. Hence, the photic regime of a bird, subjected to externally imposed light stimuli, beginning at wake-up time and varying in duration, can be thought of as comparable with the light-mediated consequences of human wakefulness, with varying times of sleep onset. If the human subject were, for any reason, to delay his sleep onset from its usual circadian timing, he would presumably be administering unusual light stimuli to his pacemaker. Should the phase-shifting effects of light stimuli on the human wake–sleep pacemaker qualitatively resemble those in the sparrow, then one might well expect the sort of relationship between phase of sleep onset and duration of sleep seen in the human subjects, arising as a consequence of repeated self-administered light stimuli.

The assumption involved here, that the wake–sleep pacemakers of human and sparrow respond to the phase-shifting effects of light in qualitatively similar ways, represents an extremely speculative leap beyond the limits of experimental data. Light has been shown to be the dominant Zeitgeber for the circadian rhythms of all nonhuman species; and because bird and human are both day-active creatures, one might well expect evolution to have produced certain similarities in their responsiveness to light, at least very crude resemblances, so as to permit entrainment of their pacemakers with comparable, ecologically appropriate phase control. There is, however, no necessity for the resemblances to go as far as assumed here; in fact I know of no unequivocal demonstration to date that light can directly produce phase shifts of any sort in the human pacemaker. On the contrary, Wever's attempts (6) to synchronize human subjects with lighting regimes suggest that the human pacemaker may be appreciably less sensitive to light than that of birds (see also ref. 4). Nevertheless, while admitting the validity of such reservations, I find this hypothesis attractive because it does not involve the interpretation (R. A. Wever, *this volume*; R. E. Kronauer *this volume*) that the human wake–sleep pacemaker is unique in having or developing an intrinsic period that is many hours longer than the values seen in circadian rhythms of all nonhuman vertebrates.

Phase shifting by light stimuli would, of course, offer only partial explanation for why human subjects occasionally show the phenomenon of internal desynchronization, with attendant wake–sleep cycles that are so far beyond the usual circadian range. One must also explain why humans during this sort of experiment would fail to "listen to" their circadian pacemakers, at the internally proposed time of sleep onset, and postpone that onset for so many hours, during

which light effects of the sort here postulated might arise. Perhaps the answer lies in the cognitive interaction of a human subject with the protocol typical of such studies in temporal isolation. In the Erling experiments (6), the subjects are instructed to follow their normal, regular daily routine, with three primary meals per day, but to avoid afternoon naps if at all possible. One of the experimental subjects who experienced internal desynchronization (6, Fig. 57, right) has retrospectively reported the following sort of subjective experience: "I'm tired now, but this is probably not my bedtime; I have only eaten breakfast and lunch, so I presume that this sleepiness represents an urge to nap, perhaps because I'm cooped up here with too little chance to exercise. I won't go to bed now, but try some gymnastics instead, to wake me up." As another example of such cognitive interactions, discussion speakers at this symposium have called attention to the fact that in other laboratories, the decision to retire "for the night" during temporal-isolation studies is a much more momentous one than simply climbing into bed; the subject must be fully instrumented (EEG, etc.) by laboratory personnel before sleep is permitted. Could this aspect of the protocol lead some subjects, some of the time, to postpone the decision to go to bed, to procrastinate, waiting to see if the inner alarm clock is really giving an irrevocable bedtime signal? We are all familiar with the experience that cognitive issues of far less direct consequence, such as good company or interesting reading material, can lead us to postpone our sleep time—though seldom to the extreme sometimes observed in temporal-isolation experiments.

Perhaps, as Winfree (*this volume*) has suggested, cognitive processes of these sorts could be responsible also for the discouraging outcome of his search for any consistent correlation between the timing of sleep onset and the prior features of a given subject's wake–sleep behavior. (Note that the simulation results shown in Fig. 1 depend primarily on the time at which the light stimulus ends, i.e., time of sleep onset, with only minor influence of light onset, i.e., preceding wake-up time.) Although there is no direct evidence yet available to indicate whether or not cognitive factors can influence the outcome of human temporal-isolation experiments, that absence could conceivably reflect only the failure systematically to collect, analyze, and interpret adequate data on the *subjective* aspects of the isolation experience.

Whether or not the preceding interpretation contains an element of the truth, the predictability of sleep duration on the basis of phasing of sleep onset is an exciting empirical result, a central feature of these remarkable experiments that demonstrates the conspicuous involvement of a circadian pacemaker even in these very noncircadian data. The speculations in this chapter about those data are offered in the spirit that alternative hypotheses are essential to the progress of experimental science. The hypotheses suggested here are sufficiently different from the others currently under consideration that a variety of critical experiments to distinguish between alternatives could be designed. One of the most obvious of these would be to attempt to induce internal desynchronization by sleep deprivation of a diurnal animal (a monkey, for example) during free-

running conditions, first in a well-lit environment and then in total darkness. If self-induced phase shifting of the wake–sleep pacemaker by lighting is responsible for human internal desynchronization, the sleep-deprived monkey should show phase shifting of its wake–sleep rhythm in a well-lit environment, resulting in internal desynchronization, but should not be so affected in darkness. I deeply regret that I am not now in a position to undertake such an experiment myself.

SUMMARY

The complex dependence of sleep duration on circadian phase of sleep onset that is seen in human experiments bears a clear resemblance to simulation results that predict the duration of sleep of a sparrow following abnormally long light stimuli. This resemblance suggests the hypothesis that the human results may be due to (a) cognitive interactions of the human subject with the experimental protocol such that he sometimes forces himself to remain awake far after the time suggested by his circadian pacemaker and (b) phase shifting of that pacemaker by self-administered light stimuli during such intervals of prolonged wakefulness. This hypothesis implies that the free-running period of the underlying wake–sleep pacemaker could be near 24 hr, even when the average period of the observed rhythm is much longer.

ACKNOWLEDGMENTS

This work resulted in part from informal conversations, particularly those with Drs. S. Daan, G. Dirlich, C. Eastman, and A. Winfree, that took place during the course of the symposium reported in this volume. Dr. Winfree has also offered valuable comments on the manuscript. This work was supported in part by Grant PCM-7719949 from the National Science Foundation.

REFERENCES

1. Czeisler, C. A., Weitzman, E. D., Moore-Ede, M. C., Zimmerman, J. C., and Knauer, R. S. (1980): Human sleep: Its duration and organization depend upon its circadian phase. *Science*, 210:1264–267.
2. Czeisler, C. A., Zimmerman, J. C., Ronda, J. M., Moore-Ede, M. C., and Weitzman, E. D. (1980): Timing of REM sleep is coupled to the circadian rhythm of body temperature in man. *Sleep*, 2:329–346.
3. Enright, J. T. (1980): *The Timing of Sleep and Wakefulness*. Springer-Verlag, Berlin.
4. Lewy, A. J., Wehr, T. A., Goodwin, F. K., Newsome, D. A., and Markey, S. P. (1980): Light suppresses melatonin secretion in humans. *Science*, 210:1267–1269.
5. Menaker, M. (1968): Extraretinal light perception in the sparrow. *Proc. Natl. Acad. Sci. USA*, 59:414–421.
6. Wever, R. (1979): *The Circadian System of Man*. Springer-Verlag, Berlin.

Mathematical Models of the Circadian Sleep-Wake Cycle, edited by M. C. Moore-Ede and C. A. Czeisler. Raven Press, New York © 1984.

General Discussion

Dr. Kronauer: Both Dr. Czeisler's data and Dr. Jouvet's data show that there is a preponderance of the short sleep periods and not very many of the long sleep periods in internally desynchronized subjects. But I look at the data that Dr. Winfree put together from Dr. Zulley's experiments and I find that in this distribution there are a lot of long sleeps and not very many short sleeps. It is a puzzle; I do not understand what might make that kind of difference.

Dr. Wever: There is a difference in the way the data are presented that maybe you do not realize. What Dr. Czeisler did was to relate sleep duration to the educed waveform of the temperature rhythm rather than to the waveform of the rhythm on the given day.

Dr. Kronauer: But that does not explain why there should be more long and fewer short sleep episodes, as opposed to what Dr. Czeisler reported.

Dr. Wever: I would say our study used much more data. [*Laughter*] The difference in the results depends on whether the educed minimum or actual minimum is used. The actual minimum is shifted away from the mean of the educed rhythm systematically.

Dr. Czeisler: Well, this analysis was done because we were relating the timing of sleep–wake processes to the circadian phase at which they were occurring and not the absolute temperature per se. There is an evoked response when you go to sleep with a lowering of temperature immediately thereafter. In contrast, when Dr. Wever subjectively determines the minimum in each cycle, that process requires a subjective decision to be made as to the time of the minimum in each curve. We wanted to use an objective method which related sleep to the phase of the cycle. So we determined the average period length by either spectral analysis or various other techniques that we have developed and then educed an average waveform around that period. In those cases, we eliminate the evoked component of the cycle and are able to look at the relationship of these processes to the phase of the circadian oscillator controlling the temperature cycle.

Dr. Schulz: Maybe one of the differences between Dr. Jouvet's data and Dr. Zulley's data was that Dr. Jouvet's single experiment was done in a cave where the temperature was rather low. In the German bunker experiments of Dr. Zulley and Dr. Wever, however, the ambient temperature was self-selected.

Dr. Czeisler: Even in the isolation facility in New York, where the ambient temperature is self-selected (along with the light–dark cycle), there is a preponderance of self-selected sleep onsets near the trough of body-temperature cycle. I think that it also may have to do with the particular subjects that Dr. Zulley selected, because the greater the length of the average period of the cycle (in other words, the fewer clusters that return to the near 25-hr period), the fewer cycles will have short sleeps. You can take sections of Dr. Jouvet's data, for example, which have a preponderance of the short sleep episodes and other sections which have a preponderance of the longer sleep episodes. But, in general, if you average many subjects together, since most of them start off at the beginning of desynchronization with a period length of around 30 hr, there is a preponderance of clusters of short sleep episodes.

Dr. Wehr: If you use the educed waveform to relate the sleep duration to the phase, and if there is a certain variation from cycle to cycle, you will have a rather imprecise definition of the actual phase at which each of those temperature cycles relates to the

sleep episode. On the other hand, Professor Wever's method, if it in fact accurately identified the specific phase of temperature oscillator, would correct for that problem. But is the actual temperature rhythm really so tightly related to the actual phase of the driving oscillator? So I do not see either method as being superior.

Dr. Czeisler: Well, I do. [*Laughter*] And the reason is twofold: First of all, if you look at the actual data from one of our long free-running experiments with up to 6 months of data, you can see that the temperature cycle does not have the kind of phase shifts which Dr. Wever presupposes. The troughs of the temperature cycle are remarkably consistent and are without shifts. In fact, it is the absence of shifts during free run that really surprises us, given the nature of the sleep–wake cycle. Second, even if there were such shifts, I do not see the method that Dr. Wever uses as any improvement over detecting the phase. In fact, his method is much more subject to the evoked responses of going to sleep and of activity which are superimposed on the endogenous circadian system.

Dr. Enright: Like Serge Daan, I consider myself an evolutionary biologist, but I think one of the most important and unanswered questions here is why we frequently see internal desynchronization in humans and do not see it in any other organism, even the primates that have been looked at. The one obvious correlation I see is that in humans there is the potential influence of cognitive factors. If you have got a good book, you are likely to stay up later, but if you are bored, you are likely to go to bed earlier. Now we hear from Art Winfree that the time of sleep onset, the characteristic that I would expect to be most readily influenced by such cognitive factors, is apparently unpredictable with respect to other aspects of the rhythm. Can anybody offer speculation about why we do not see internal desynchronization in other animals?

Dr. Moore-Ede: The contrasts between the human and the animal studies are fraught with more problems than just differences in cognitive activity. First, Dr. Weitzman and Dr. Czeisler have pointed out that there is little obvious evidence that subjects are staying awake in order to finish some task whenever they are displaying a very long period of wakefulness and, therefore, extending their time of sleep onset. Second, the experiments that have been done in monkeys have shown spontaneous internal desynchronization by all criteria between other variables such as the urinary rhythms and the behavioral rhythms, but not so far like the human subjects, between body-temperature rhythms and activity rhythms. Now the work of Dr. Gander in my lab reported at this meeting shows that the activity rhythm shows a rather greater variability than the temperature rhythm and may, in fact, show quite large dissociations in phase between these two rhythms in free-running monkeys. We have seen the data of Dr. Wirz-Justice, which suggest such phenomena may exist in rodents treated with anti-depressants. So, given the correct conditions, we very well may be able to show internal desynchronization in other species. It is quite possible that playing around with various agents, either that change the period or by using self-selected light–dark cycles, we may be able to induce internal desynchronization.

Dr. Czeisler: I think that it is interesting to note that it is the people who have not done actual experiments with human subjects who are most concerned about the volitional aspect and that those of us who have actually done those experiments and have seen that volitional factors do not play a part in internal desynchronization are least concerned.

Dr. Enright: In animal experiments you do not see the noncircadian periods you see so often in humans.

Dr. Kripke: In the human data the idea of a single 24-hr sleep–wake cycle is reasonable because many people will stay awake for 16 hr and go to sleep for 8 hr. Animals do not do that. A rat, for example, wakes up every hour, at the least, for a period of time. So the human is relatively unique compared to all our animal models, which have polycyclic sleep.

Mr. Pilato: I've heard a quite narrow conception of what cognitive activity is here, and that is reading a book late at night.[*Laughter*] I would like to refer to an anecdote that was reported by a newspaper reporter, who took part in one of the experiments in the Montefiore lab, which I thought was very interesting. She said that after a while her normal sequence of daily activity started to break down, and her normal timing mechanisms within the day were harder to determine because she could not tell the time by the fact that she had just finished one typical daily activity followed by another. So I would like to propose that humans, under normal circumstances, utilize not only light and dark to time themselves, but also their sequence of activities as well as external cues such as the position of the sun or the use of a wristwatch. In the absence of such cues, humans are at a loss because they depend on those kinds of nonnatural environmental cues.

Dr. Czeisler: I would also like to point out that it is not just the last part of the day which is extended. It is not just that they stay up late performing some activity, but the entire day changes its character under these conditions. The time between waking up and having breakfast might be 6 or 8 hr; the time between breakfast and lunch might be 8 to 10 hr. So dinner usually comes just before the subjects go to sleep, even when they are on a 40- or 50-hr day. In the outside world we might have gone through breakfast, lunch, and dinner before the subject even has breakfast, despite the fact that he got up at the same time. So it is not just an isolated part of the day that has lengthened under this circumstance, but the subject's entire perception of the length of his day.

Dr. Edmunds: I would like to challenge the conclusion that internal desynchronization has not been observed in any organisms besides humans. A couple of years ago I published data suggesting that the rhythms of photosynthetic capacity can show dissociated circadian rhythms in free-running *Euglena* over an interval of 2 or 3 days, with virtually any phase relation observable, suggesting that the detection of internal desynchronization may depend on what you look at.

Dr. Weitzman: I would like to expand on what Dr. Czeisler said. If you ask a subject who has been up for 30 some hours if anything is different (being careful not to imply to him that something might be different), almost invariably you get a statement, "Well, maybe the day seemed a little longer but basically it was unchanged." On a 50-hr day, you often will find that they feel like they are ready to go to sleep but they have not had dinner yet. And so they are conflicted—"Should I have dinner, should I go to sleep?" In fact, sleep generally takes precedence, and they say, "I just cannot stay awake anymore," and they will go without dinner. Thus, the whole structure of the day, the whole internal organization, changes, but the subject is totally unaware of the time that has elapsed in these very long days.

Dr. Wever: We find there is a very strong correlation between internal desynchronization and the age of the subject. We can divide our subjects into two groups: a younger group from 17 to 35 or 40, only a few of whom show internal desynchronization. In the elder subjects, up to age 73, 80% show internal desynchronization.

Dr. Weitzman: We have not seen such a high incidence of internal desynchronization in older subjects. We have studied a series of older subjects up to the age of 84, and there have been some who have desynchronized, but most of them have not. We do not have a very large series, but it suggests to us that it is more the length of time in isolation rather than the age of the subject that is important.

Dr. Winfree: Dr. Weitzman, on these 50-hr days, do your subjects eat 6,000 calories per cycle?

Dr. Weitzman: No. They do not.

Dr. Winfree: If they eat three meals a day, do they get skinny after a while?

Dr. Weitzman: Yes, they do. In fact, some subjects increase the amount per meal; they will say they feel hungry. But another subject might say, "I do not know what you are doing here, but you are not giving me enough food to eat. You weigh me every day, and I am hungry all the time and I am losing weight every day." He was only eating three meals and a snack on each 50-hr day. So we told him, "If you want more food or more meals, just ask for it; you have that choice. You can do anything you want." That's our paradigm: We give them any amount of food at any time; it is a wonderful hotel! But he did not choose to do that. He continued to eat essentially three or four meals a day and just increased the size of the meal somewhat, and thus continued to lose weight during that time. It is a good weight-losing technique, a little expensive, but. . . . [*Laughter*]

Dr. Lewy: Why does internal desynchronization occur either exclusively or at least more often in humans as compared to other species? Second, why does internal desynchronization occur when it does occur in the free-running studies? I think one factor that we should consider is the light–dark cycle and the sensitivity of humans to light. Humans require much brighter light for the same suppression of melatonin than other species. Another finding that we have not published yet is that there seems to be quite a large variability in light sensitivity in individuals, and there may be a seasonal rhythm in light sensitivity in man. During the summer our subjects are relatively insensitive to light, as compared to when they are studied in the winter. So I would suggest to those of you who have data that you look at what time of the year the experiments were conducted.

Dr. Borbély: Two basically different models of the sleep–wake process have been proposed. One of them postulates a single oscillator (Dr. Daan's and Dr. Eastman's model), and the other has two oscillators (Dr. Kronauer's model). It is crucial that we think of experiments which could determine which model is more valid. As Dr. Serge Daan has proposed, sleep deprivation is a method which has not yet been sufficiently exploited. It seems to me that along these lines a critical experiment could be designed.

Subject Index