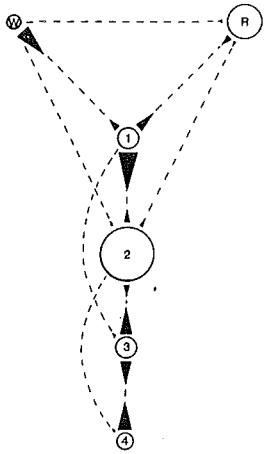
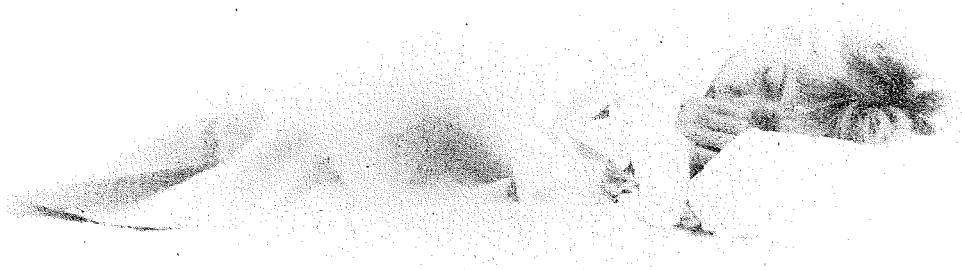
MODEL-BASED MONITORING

OF

HUMAN SLEEP STAGES





Bob Kemp

MODEL-BASED MONITORING OF HUMAN SLEEP STAGES

P r o e f s c h r i f t

ter verkrijging van de graad van doctor
aan de Universiteit Twente,
op gezag van de rector magnificus,
Prof. dr. ir. H.H. van den Kroonenberg,
volgens besluit van het College van Dekanen
in het openbaar te verdedigen
op donderdag 11 juni 1987 te 16.00 uur

door

Bastiaan Kemp

geboren op 2 augustus 1951 te Ede

Promotoren:	Prof.dr. H.A.C. Kamphuisen,	Rijks Universiteit Leiden
	Prof.ir. E.W. Gröneveld,	Universiteit Twente
	Prof.dr. F.H. Lopes da Silva,	Universiteit van Amsterdam
Overige leden der promotiecommissie:		
	Prof.dr.ir. J.H. van Bemmel,	Vrije Universiteit Amsterdam
	Dr. D. Kleima,	Universiteit Twente
	Prof.dr.ir. H. Kwakernaak,	Universiteit Twente
	Prof.dr. S.L.H. Notermans,	Katholieke Universiteit Nijmegen

Dit onderzoek is uitgevoerd aan het Academisch Ziekenhuis Leiden, afdeling Klinische Neurofysiologie, NL-2333 AA Leiden, Nederland.

Omslag: G.J. van de Giessen / B. Kemp
Druk : Krips Repro, Meppel.

CIP-DATA KONINKLIJKE BIBLIOTHEEK DEN HAAG
Kemp, Bastiaan
Model-based monitoring of human sleep stages / Bastiaan (Bob) Kemp.
- [S.l. : s.n.]. - Ill.
Thesis Enschede. - With ref. - With summary in Dutch.
ISBN 90-9001448-9
SISO 415.5 UDC 159.963.001.5
Subject heading: human sleep ; model-based monitoring.

CONTENTS

Preface

Although the major part of this research started in 1984, my interest in simple models of rhythmical brain activity, based on white noise, feedback loops and Markov chains, dates back to 1975-1977. At that time, I was working both as a volunteer with Professor Dr. F.H. Lopes da Silva (now at the University of Amsterdam) at the TNO Institute of Medical Physics, Utrecht, and on my masters thesis under the supervision of Professor Ir. E.W. Gröneveld at Twente University, Enschede. During that period and ever since, I have received regular support and criticism from them both.

The Wonham-filter (ch.III) was suggested to me in 1976 by Professor Dr. Ir. H. Kwakernaak from Twente University. However, at that time our mathematical models were too complex to allow application of the filter and we did not yet realize that the existence of feedback loops in the brain would enable us to make the required simplifications. These were first exploited in a 1978 masters thesis in Twente by Ir. H.A.P. Blom. (now at National Aerospace Laboratory in Amsterdam). It took a course (in 1980) on stochastic filtering and control by (a.o.) Dr. Ir. J.H. Van Schuppen of the Centre for Mathematics and Computer Science in Amsterdam, followed by many discussions about possible applications (ch.III and IV), to rediscover the Wonham filter. At the end of that year, I rederived the results obtained by Blom, but this time using both this filter and a still somewhat more simplified model. Although stochastic filtering theory methods were abandoned in 1985 (ch.V), they taught us how to deal properly with the physically and mathematically troublesome white noise in our models.

Ever since my appointment in 1977 as a medical physicist (although I am not a physicist) in the Department of Clinical Neurophysiology in Leiden, Professor Dr. H.A.C. Kamphuisen has made me feel at home in this clinical environment and he gave me the opportunity to initiate new research direc-

tions. As head of the physics and informatics group and as a staff member, I have always tried to find an efficient balance between my own research and other projects in the department. Still, the innovating and reliable contributions of the group (Drs. J.M. Franzen, A.J.M.W. Janssen, ing. C.G.S. Kramer and many passing guests) have been indispensable for most of the applications in this thesis. Working with a small team in close collaboration with the medics (in particular Drs. R.J.H.M. Arts and Drs. J.G. Van Dijk) and paramedics (a.o. Mrs. G.J.M. Keulen, Mrs. M. Van Rijnsbergen and Mrs. H.D.J. Geurts) in a clinical environment has been most helpful in designing methods to improve patient care. The many problems arising from the incompatibility between biology and technology and from all kinds of human factors were in general recognized at early stages of the projects.

Dr. J.H.M. De Groen (now at the University of Limburg) and Ir. A. Kumar (University of Amsterdam) introduced me to sleep and its analysis in 1977. However, it was only after years of modelling and analysis of the waking brain, and an exciting excursion into radar quantification of neurological body movement disorders (with Dr. O.J.S. Buruma and Dr. R.A.C. Roos), that we decided in 1983 to focus on sleep, in particular on automatic sleep monitoring. I still had to find out how to deal with multichannel, multirhythmical and multistate brain activity, eye movements, muscle tone, 'future' information and how to combine these. While doing this, I had the invaluable support and cooperation of Ir. P. Jaspers (Delft University of Technology, now at Philips Medical Systems), Dr. A.N. Nicholson (R.A.F. Institute of Aviation Medicine, U.K.) and Dr. E.L.E.M. Bollen. Soon, it became clear that the sleep project would continue to grow and many sleep recordings would have to be carried out. Therefore, I also started the modifications of the portable home recorder, that have been evaluated first by Mrs. G.T.M. Van Dok in 1985.

Many technicians and volunteers gave me the opportunity to test all sorts of ideas, some of which led to this thesis and others have been placed in the custody of Mr. J. Bink (archives). Literature was traced by Mrs. A. Van Zuyden. Photographic reproductions were made by Mr. G.J. Van De Giessen and Mr. A.J. Rhijnsburger. Mrs. T. Offers and Mrs. T. Rodenburg forced our reluctant text prOCEssor to print formulae. This was only possible at the cost of a rather small printing fount, for which I apologize. Mrs. B. Vollers and Ms. M.E. Van Hoogevest respectively, language-corrected the English and provided the French texts. Because chapters II, III and IV have been, and V is intended to be, published as separate articles, they each contain more specific acknowledgements.

I hope we can continue our collaboration in unravelling some of the mysteries of sleep and sleep disorders.

I. INTRODUCTION

In this chapter we state the reasons for our interest in sleep monitoring and its automation. The major propositions, purpose and limitations of this thesis will be indicated. Because chapters II, III and IV have been, and V is intended to be, published as separate articles, they each contain their own specific introduction. Some general literature on sleep and its analysis has been listed at the end of this chapter.

Most people spend about 30% of their lifetime asleep. The urge to sleep may be so irresistable that one may even fall asleep while driving a car. Even when subjected to highly intensive, stimulating programmes, everyone will fall asleep within a few days of sleep deprivation, before any serious harm seems to have been done. Apparently, sleep must fulfill some vital function, that requires and compensates for these long, vulnerable and unproductive periods of paralysis and unconsciousness. There has been much speculation about the nature of such a function (physical or psychological restoration, energy saving, learning, etc.) but none has been confirmed. Still, about 10% of middle-aged and 20% of older people feel that they do not sleep well. These complaints may have various causes with different effects on sleep and daily performance. Sleep monitoring already makes it possible to diagnose and treat several sleep-related disorders (e.g. narcolepsy, circadian rhythm disturbances, sleep apnoea, depression, insomnia, excessive daytime sleepiness). However, as we still know very little about the physiological mechanisms of sleep, so we can hardly expect to understand its disturbances. If we wish to know more about sleep physiology, function and disorders, we have to perform thorough studies by monitoring sleep in many subjects over long periods of time.

To do this, we must specify what we mean by 'sleep'. During normal sleep, several behavioural (e.g. depth of consciousness, reactivity, body

posture), psychological (e.g. probability of recalling a dream, dream contents) and physiological (e.g. brain electrical activity, muscle tone, heart rate, eye movements, twitches of the limbs, pupil contraction, body temperature, release of growth hormone) variables are modulated in a more or less synchronous way. Several triggers (e.g. night shift, sleep deprivation, pregnancy, jet lag) and most sleep-related disorders affect the time course of this modulation, while the correlation between the variables is preserved. The same holds for interindividual differences. Some adults feel all right and perform adequately after only 4 hours sleep, while others need 10. Elderly people spend much less time asleep (particularly 'dream-sleep') than children. However, the sleeps of different individuals can still be compared, because they all show simultaneous variations of the same biological variables. This is also partly true for animals, enabling phylogenetic and animal experiment studies to be carried out. We therefore conform to the often applied operant definition of sleep as a process characterized by synchronous variations of a specific set of biological variables. In particular, possible disconnection (also called dissociation, e.g. chapter VI.2 and VI.3) of these variables will not be considered here.

Most applications, also in this thesis, are restricted to electroencephalographic (EEG), electrooculographic (EOG) and electromyographic (EMG) variables that can be recorded by means of surface electrodes on the scalp, near the eyes and on chin muscles, respectively. These recordings can be obtained at relatively low cost, little discomfort and with little risk of disrupting the sleep process. They show quite strong interrelations. The restriction is not essential: the theoretical framework in this thesis is of sufficient generality to include many more variables.

Based on the above proposition, sleep can be monitored by recording EEG, EOG and EMG. However, these signals are also subject to unpredictable

variations which are not related to sleep. A certain amount of noise filtering is, therefore, necessary. This is usually done by training EEG-technicians and neurophysiologists in the rather complex process of recognizing sleep-related signal characteristics. They perform subsequent averaging over a period of 30 seconds. Smoothing over intervals up to 3 minutes may be necessary if the 'signal-to-noise-ratio' is still insufficient. The final result is a classification of sleep into a limited number of stages (fig.1). This sleep-pattern, or hypnogram, is accepted worldwide as a basis for the study of sleep and the diagnosis of sleep disorders. Despite the standardization of terminology and classification rules, and despite intensive training programmes, different trained specialists (and even the same specialist on different occasions) produce quite different classifications. Moreover, the procedure is very laborious; clinical recording and analysis of one night takes about 10 and 5 hours, respectively. Often, several nights per subject have to be classified by more than one specialist and the resulting costs are a major obstacle in sleep research and patient care. In contrast to our neighbouring countries and the U.S.A., The Netherlands does not have a specialized sleep centre.

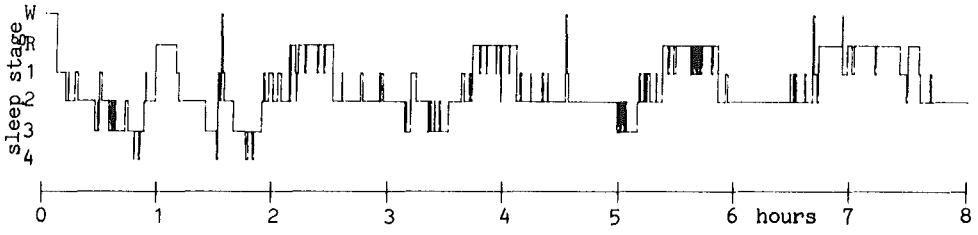


Figure 1: A typical whole-night classification of sleep in the stages Wakefulness, Rem (i.e. 'dream-sleep'), and stages 1, 2, 3 and 4 that range from drowsiness to very deep sleep.

Automation of both recording and analysis would reduce costs considerably and improve classification consistency. For a few years now, automatic recording with the subject sleeping at home in his or her own bed has been applied in several laboratories. The particular method developed by us has been operational since 1985 (ch.VI.4).

Automatic sleep analysis has been a subject of interest since the fifties. The methods were first based on analogue processing, later also on digital or hybrid computers. They mimic the human classification procedure in that they consist of preprocessors that enhance sleep-related signal components, followed by some smoother in order to remove the remaining noise. Many methods have been suggested (ch.V). Only a few have been compared to the standardized human classification, resulting in man-machine agreements ranging from 65% to 86%. The corresponding man-man agreement was established in only two studies (Lacroix and Stanus 1985, Martin et al 1972), being 8% better in both cases. These comparisons have been based on only a few nights of sleep in (except for one patient) healthy young male subjects. No system has yet been validated such that it can be relied upon in routine clinical sleep studies. All methods are based on tuning and retuning during years of experience. Most of them have not been reported in sufficient detail to enable, specify and reproduce applications. All methods combine several intermediate decisions about isolated sleep-related signal components (e.g. 'many' or 'few' rapid eye movements). Because the reliability of these decisions is not taken into account, the final decision on the sleep stage is suboptimal. Not one of the methods exploits the fact that the optimal way to obtain information about an unknown process from noisy observations, can be derived from a model that describes the statistics of both the process and the observations.

However, the physiological and medical literature now provides enough

material to develop simple stochastic models of sleep and the sleep-related observations (chapters II,III,IV,V). Based on information theory, we can then create the corresponding, optimal way to process these observations in order to obtain the desired information about sleep (ch.III,IV,V). The models introduce physiological and statistical knowledge into the sleep monitor and so we may hope that the results are closer to 'real' sleep than is possible by mimicking the, probably suboptimal, human analysis procedure. This thesis shows that the advantages of this approach were strong enough for the rapid development of a sleep monitor that performs adequately and is sufficiently simple to specify the method and interpret the results.

Literature

Association of Sleep Disorders Centers and the Association for the Psychophysiological Study of Sleep. Diagnostic Classification of Sleep and Arousal Disorders. Sleep 1979, 2:5-129.

Borbély A, Valatx JL. (eds.) Sleep Mechanisms. Springer Verlag, Berlin, 1984.

Dement WC. Some must watch while some must sleep. The Stanford Alumni Association, Stanford, 1974.

Drucker-Colin R, Shkurovich M, Sterman MB (eds). The functions of sleep. Academic Press, New York, San Francisco, London, 1979.

Guilleminault C, Lugaresi E (eds). Sleep/wake disorders. Raven Press, New York, 1983.

Hasan J. Differentiation of normal and disturbed sleep by automatic analysis. Acta Physiol Scand, supplementum 1983, 526:1-103.

Lacroix B, Stanus E. New algorithms for on-line automatic sleep scoring, and their application to mini and micro-computer. Journal A 1985, 26:91-97.

Lairy GC. Critical survey of sleep stages. In: Proc 3rd Europ Congr Sleep

Res 1976:170-184. Karger, Basel, 1977.

Martin WB, Johnson LC, Viglione SS, Naitoh P, Joseph RD, Moses JD. Pattern recognition of EEG-EOG as a technique for all-night sleep stage scoring. Electroenceph clin Neurophysiol 1972, 32:417-427.

McGinty DJ, Drucker-Colin R, Morrison A, Parmeggiani PL (eds). Brain Mechanisms of Sleep. Raven Press, New York, 1985.

Miles LE, Dement WC. Sleep and Aging (1-8). Sleep 1980, 3:119-220.

Monroe JL. Inter-rater reliability and the role of experience in scoring EEG sleep records: phase I. Psychophysiology 1969, 5:376-384.

Rechtschaffen A, Kales A (eds). A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Public Health Service, US Government Printing Office, Washington DC, 1968.

Spiegel R. Sleep and sleeplessness in advanced age. MTP Press Ltd Lancaster Spectrum Publ Inc, New York, 1981.

Weitzman ED (ed). Advances in Sleep Research I. Spectrum, New York 1974.

II. SIMULATION OF HUMAN HYPNOGRAMS USING A MARKOV CHAIN MODEL

Bob Kemp and Hilbert A.C. Kamphuisen

Reproduced from Sleep 1986, 9:405-414

Summary

A Markov chain model has been proposed as a mechanism that generates human sleep stages. A method for estimating the parameters of the model, i.e. the transition probabilities (rates) between sleep stages, has been introduced and applied to 95 hypnograms taken from 23 subjects. The rates characterize interindividual differences and nightly variations of the sleep mechanism, related to sleep-onset behaviour, to the decreasing amount of slow wave sleep in the course of the night, and to the REM-NREM periodicity. The model simulates both probabilistic and the above mentioned predictable dynamics of sleep, but only if these time-varying, individual rates are applied.

II.1. Introduction

In the course of a night, human sleep seems to travel through various stages, following a rather unpredictable pattern (fig.1). Despite the widely applied classification of sleep into a limited number of discrete stages (e.g. Rechtschaffen and Kales 1968), the precise definition and the functional significance of these stages is not clear and other stages probably also exist (Rechtschaffen and Kales 1968, Lairy 1977, Broughton 1982, Parmeggiani et al 1985). However, the classification is based upon clear electrophysiological events that occur during sleep and many investigators have demonstrated the correlation between these stages and not only various somatic, autonomic and biochemical sleep-related phenomena (e.g. Parmeggiani et al 1985), but also pathological aspects of sleep (e.g. Broughton 1982,

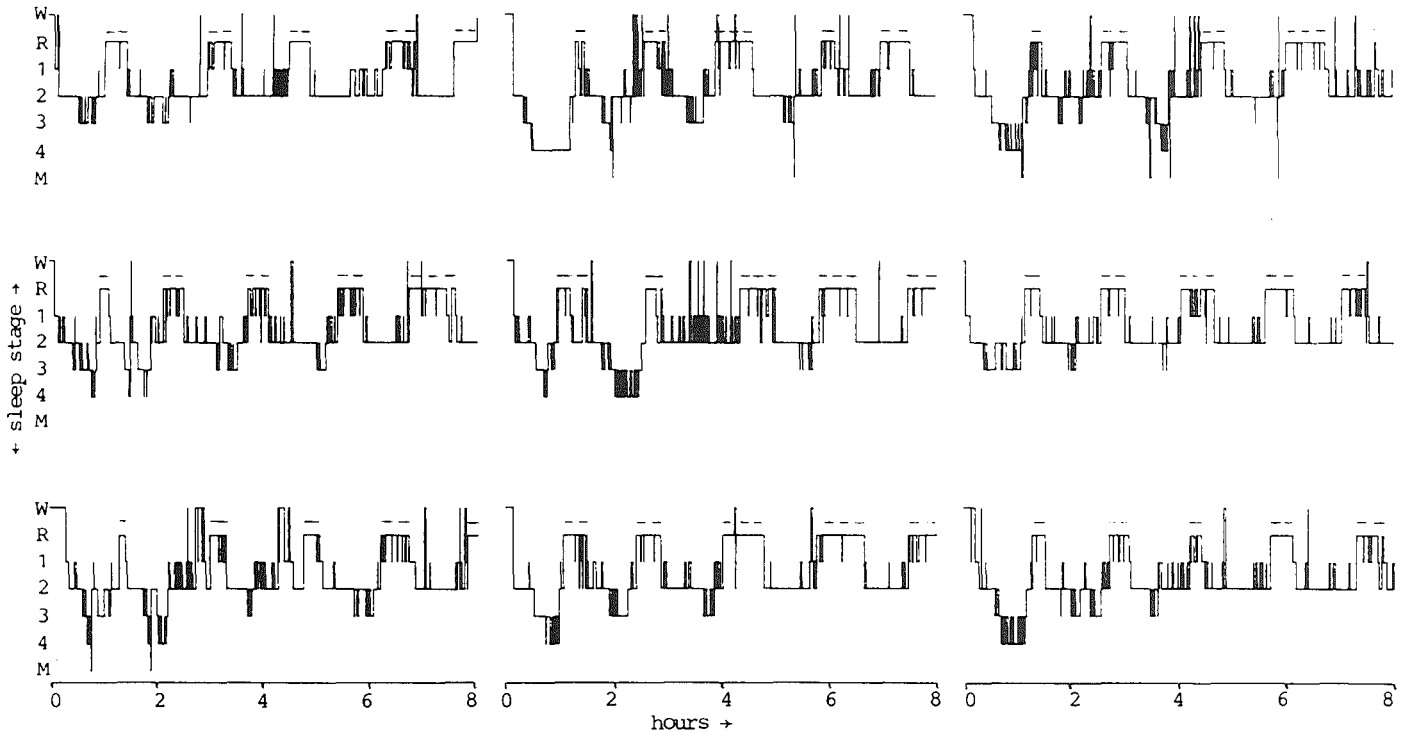


Figure 1: Hypnograms of nine nights of subject AD. Lights-off at 0 hour. Sleep stages Wakefulness, Rem, stage 1, stage 2, stage 3, stage 4 and Movement time. Note the 90 minute REM-NREM period and the decreasing amount of stages 3 and 4 in the course of the nights. Dashes: REM-'blocks'.

Dement and Kleitman 1957, Feinberg et al 1967, Miles and Dement 1980, Spiegel 1981). Therefore, it may be useful to develop a model that can simulate hypnograms, since this may suggest a sleep mechanism, parameters to characterize sleep (Hermann and Kubicki 1984) and methods to analyse sleep (c.f. this paper and also Kemp et al 1985).

No comprehensive model of a hypnogram-generating mechanism is available, but several models describe various aspects of it. A physiological model of the REM-NREM sleep cycle mechanism in the cat (Hobson et al 1975, McCarley and Hobson 1975, McCarley 1980) can also be used to simulate various human sleep characteristics (Beersma et al 1984). Several phenomenological models have been based upon reproducible, predictable aspects of human hypnograms. There are two categories of such models. (**a**) Deterministic models use time-varying functions to describe systematically occurring characteristics like the REM-NREM periodicity, the decreasing amount of slow-wave sleep (stages 3 and 4) in the course of the night and the relationship between sleep and the circadian rhythm (Beersma et al 1984, Borbély 1982, Winfree 1982, Lawder 1984, Borbély 1984, Daan et al 1984). (**b**) Probabilistic models use stochastic processes to describe statistical properties of hypnograms like the variability of the REM-NREM cycle, the short interruptions within REM 'blocks' (fig.1), and the random (in time and in direction) transitions between stages (Zung et al 1965, Yang and Hursch 1973, Bowe and Anders 1979, Ursin et al 1983).

Although some stochastic variation has been introduced into one of the deterministic models (Daan et al 1984) and some time-varying functions into probabilistic models (Yang and Hursch 1973, Ursin et al 1983) the resulting mechanisms cannot, and were not intended to, simulate all probabilistic and deterministic aspects of hypnograms. To our knowledge, no simulated hypnograms have been reported.

In this article we present a simple model that contains all the above-mentioned probabilistic and deterministic aspects of sleep. The model will be introduced in sufficient detail to enable, specify and reproduce applications. We investigated its ability to characterize the nightly dynamics and the individuality of the sleep mechanism. Finally we will show some simulated hypnograms.

II.2. Model

We will conform here to the standardized form of hypnograms (Rechtschaffen and Kales 1968) in which sleep is classified into the set of seven stages wakefulness, REM, stage 1, stage 2, stage 3, stage 4 and movement time, abbreviated W, R, 1, 2, 3, 4 and M, respectively (fig.1).

Transitions between these stages occur in unpredictable directions and at unpredictable moments. However, they seem to obey a probability law in which the sojourn times (i.e. the intervals between two successive transitions, during which sleep remains in the same stage) have an approximately exponential distribution (Kemp et al 1985, Yang and Hursch 1973, Williams et al 1964, Bŕezinová 1975). This observation suggests that the transitions may be generated by a continuous-time Markov chain process (e.g. Larson and Shubert 1979:ch.3.3). Such a model was proposed briefly by Zung et al in 1965. Yang and Hursch (1973) concluded that the model is inadequate because the sojourn times in a group of individuals do not fit geometric (the discrete-time analogue of continuous-time exponential) distributions. However, different individuals show different sleep characteristics (e.g. Spiegel 1981), that are also reflected by the model (as set forth in this article). Because of this and because the average of different (individual) geometric distributions is not geometric, the argument of Yang and Hursch (1973) does not hold. Based on this argument, a semi-Markov model has been proposed (Yang and

Hursch 1973), in which the sojourn times may have any (non exponential) distribution function. These are two-dimensional functions, both of the clock time and of the sojourn time. Specification of the model is therefore almost impossible, and in all applications (Yang and Hursch 1973, Bowe and Anders 1979, Anders and Keener 1985, Anders et al 1985) only mean sojourn times (a function of clock time only) were considered. Because these are implicitly specified in the Markov model (Larson and Shubert 1979), all results obtained by these semi-Markov parameters can also be obtained by the Markov parameters. For the same reason, the attempt to validate the semi-Markov model (Yang and Hursch 1973), in fact, supports the Markov model. A continuous-time model is preferred because real sleep is not segmented into epochs (e.g. of 30s), although hypnograms generally are.

The mechanism of the continuous-time Markov model is rather simple. If sleep, h(t), at time, t, is in a certain stage, i, then there is for each $j \neq i$ a probability, $P\{h(t+\Delta)=j \mid h(t)=i\}$, that it will be in stage j after a time interval, Δ. The Markov property implies that this probability does not depend on sleep history. Because more than one transition may occur in one interval, Δ, one usually specifies the transition rates, $a_{j|i}(t)$, over infinitely small intervals (which cannot contain more than one transition):

$$a_{j|i}(t) = \lim_{\Delta \to 0} [P\{h(t+\Delta)=j \mid h(t)=i\}/\Delta] \qquad (1)$$

This means that $a_{j|i}(t) \cdot \Delta$ is the probability that sleep jumps from i to j in the small interval, Δ. These transition rates specify completely the process that simulates sleep from wakefulness to the end of the night. The average sojourn times can be derived from the rates. In the section Simulations, the process will be discussed in greater detail.

Time dependent transition rates have been applied earlier to explain

and to quantify the apparent, but variable, ultradian periodicities in the sleep of rats (Bari et al 1981) and cats (Ursin et al 1983).

II.3. Estimation of transition rates

The parameters of the model, i.e. the transition rates, have been estimated from 95 hypnograms taken from 23 healthy male volunteers, aged 18 to 30 years, with previous experience as subjects for polygraphic sleep recordings. Records were scored by two analysts according to the Rechtschaffen and Kales criteria (see Acknowledgement).

The maximum likelihood estimator for the transition rate, $a_{j|i}(t)$, reads (Snyder 1975:ch.2.4):

$$\hat{a}_{j|i}(t,\Delta) = n_{j|i}(t,\Delta)/T_i(t,\Delta) \qquad (2)$$

where $T_i(t,\Delta)$ is the total time spent in stage i and $n_{j|i}(t,\Delta)$ is the number of transitions from stage i to stage j. Both $T_i(t,\Delta)$ and $n_{j|i}(t,\Delta)$ are counted in an interval $[t-\Delta/2,t+\Delta/2]$ of duration Δ around time t. The computation of $\hat{a}_{j|i}(t,\Delta)$ from more than one hypnogram, called averaging in the sequel, is performed by counting $T_i(t,\Delta)$ and $n_{j|i}(t,\Delta)$ from all the corresponding intervals around t. In the following, the arguments (t) and (t,Δ) will be omitted for brevity. Within small (relative to the nightly variations of the sleep mechanism, e.g. the REM-NREM periodicity) intervals, Δ, the process is assumed to be homogeneous, i.e. the transition rates are constant. In that case, the estimator is unbiased and its variance depends on T_i and $a_{j|i}$ as follows (Snyder 1975:ex.2.4.8).

$$\mathrm{var}\{\hat{a}_{j|i}\} = a_{j|i}/T_i \qquad (3)$$

A 70% confidence interval for $a_{j|i}$ can be approximated by ±1 standard deviation, i.e.:

$$-\sqrt{a_{j|i}/T_i} \leq a_{j|i}-\hat{a}_{j|i} \leq \sqrt{a_{j|i}/T_i} \qquad (4)$$

This can be transformed into two quadratic inequalities in $\hat{a}_{j|i}$ and $a_{j|i}$ (Fleiss 1981:ch.1.4). Their solutions express the confidence interval as a function of $n_{j|i}$ and T_i:

$$\hat{a}_{j|i} + \{½-\sqrt{\hat{a}_{j|i}\cdot T_i+¼}\}/T_i \leq a_{j|i} \leq \hat{a}_{j|i} + \{½+\sqrt{\hat{a}_{j|i}\cdot T_i+¼}\}/T_i \qquad (5)$$

or, equivalently:

$$\{n_{j|i} + ½ - \sqrt{n_{j|i}+¼}\}/T_i \leq a_{j|i} \leq \{n_{j|i} + ½ + \sqrt{n_{j|i}+¼}\}/T_i \qquad (6)$$

If $n_{j|i}$=10, the size of this 70% confidence interval is 60% of $\hat{a}_{j|i}$. If $n_{j|i}$=100, it is 20% of $\hat{a}_{j|i}$.

i \ j	W	R	1	2	3	4	M
W	-	.000149	.007771	.000130	.000000	.000000	.000000
R	.000221	-	.001409	.000338	.000000	.000003	.000003
1	.001363	.003211	-	.011243	.000000	.000000	.000000
2	.000249	.000405	.001069	-	.001033	.000000	.000021
3	.000137	.000026	.000231	.005195	-	.003734	.000128
4	.000028	.000000	.000028	.000198	.005777	-	.000156
M	.000000	.000000	.013492	.017460	.001587	.000000	-

Table 1: Average whole night transition rate estimates, $\hat{a}_{j|i}$ in s^{-1} (i.e. average number of transitions-to-stage-j per second-in-stage-i). See Model section for definition of abbreviations.

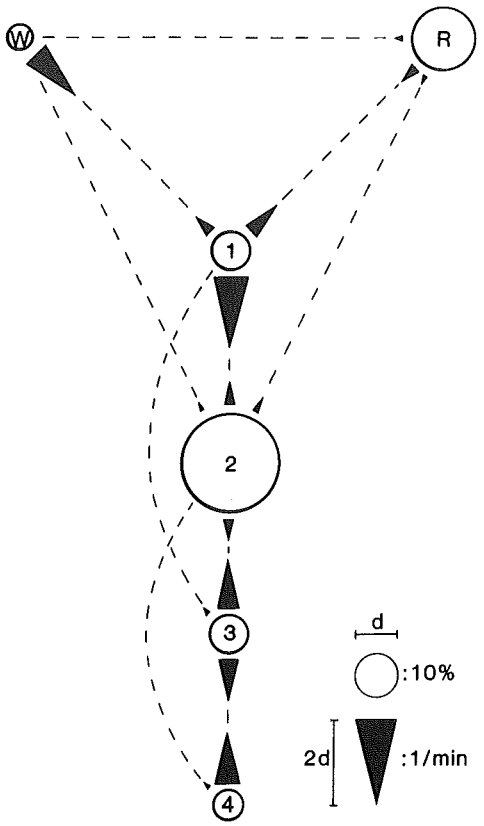


Figure 2: Average whole-night sleep structure as estimated from 46 hypnograms of 23 subjects. Circle areas are proportional to the percentage of time spent in the corresponding stage. Arrows indicate the directions of possible stage transitions. Arrow areas are proportional to the corresponding transition probabilities, i.e. rates (from table 1). The calibrations (lower right) of circles and arrows have areas $\pi d^2/4$ and d^2, which correspond to 10% and 1/min, respectively. Stage M (<0.1%) and some very unlikely transition possibilities (<1/100min) are too small to be reproduced. Note, for example, that in stage 1, the most likely transition is to stage 2.

A first impression of the absolute value of the rates has been obtained by estimating average whole-night transition rates (table 1 and fig.2) over an interval, Δ, spanning the first 8 hours of the hypnogram and by averaging over 46 hypnograms (23 subjects, 2 nights each). Although figure 2 allows a rapid inspection of some important sleep characteristics, it is biased because it neglects changes of the sleep mechanism that occur in the course of the night. Therefore, we investigated the time course of the transition rates of which the whole-night estimates were based upon at least 120 transitions (i.e. $n_{j|i} \geq 120$). The 8 hour interval has been divided into 32 intervals, Δ, of 15 minutes, each yielding an estimate if $T_i > 0$. Averaging has been

performed over the same 46 hypnograms. Highly significant ($p<0.0001$, i.e. more than 20 of the 32 confidence intervals did not include the whole-night average) inhomogeneities were present in 5 rates. In these cases we proposed a subjectively smoothed (by hand) time course (fig.3). Besides some simple trends that seem to be related to sleep-onset and slow-wave sleep, figure 3 shows periodicities in $\hat{a}_{R|2}$, $\hat{a}_{3|2}$ and $\hat{a}_{R|1}$ that may account for the REM-NREM periodicity.

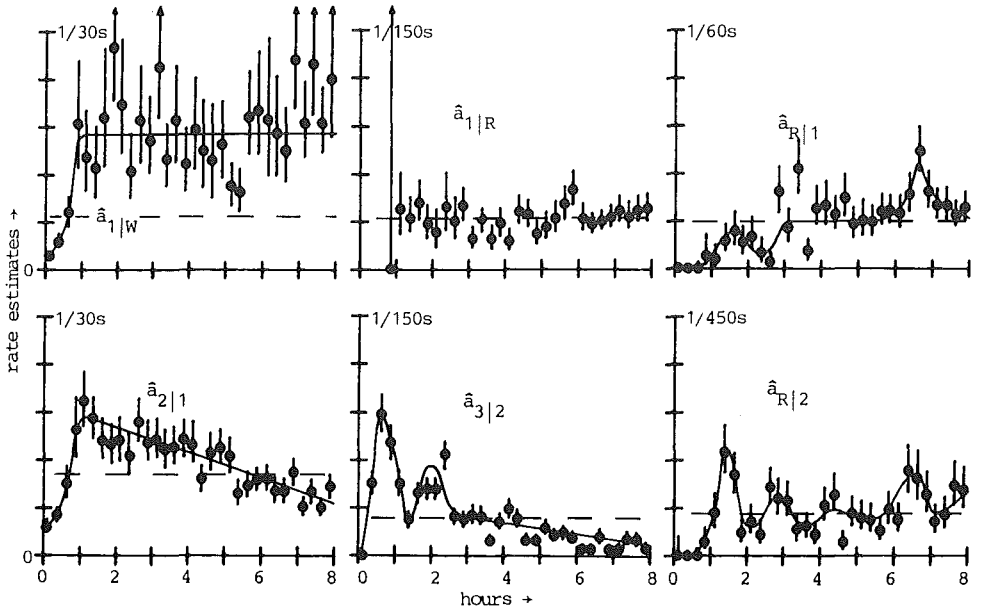


Figure 3: Nightly variations of average (23 subjects) transition rates. Circles, estimated rate, $\hat{a}_{j|i}$, from stage i to stage j. Vertical bars, approximate 70% confidence interval. Dashes, whole-night average of table 1 and figure 2. Solid line, smoothed (by hand) time course. Note the increasing tendency to fall asleep in $\hat{a}_{1|W}$ and $\hat{a}_{2|1}$ in the first hour and the periodicities in $\hat{a}_{R|1}$, $\hat{a}_{R|2}$ and $\hat{a}_{3|2}$ in the first 3 hours. Note the decreasing tendency, $\hat{a}_{3|2}$, to reach stage 3 in the second half of the night. $\hat{a}_{1|R}$ serves as an example of a very constant rate.

Because such periodicities may be obscured by interindividual differences, we have estimated these time courses separately for the 8 subjects from

whom at least six (maximum 10) hypnograms were available. Only within-subject averaging has been performed over these (six to 10) hypnograms. These individual estimates indeed show more pronounced periodicities, especially in subjects ND, AD and NS (fig.4). The three rates seem to be synchronized, $a_{3|2}$ being opposite to $a_{R|2}$ and $a_{R|1}$. They show clear interindividual differences.

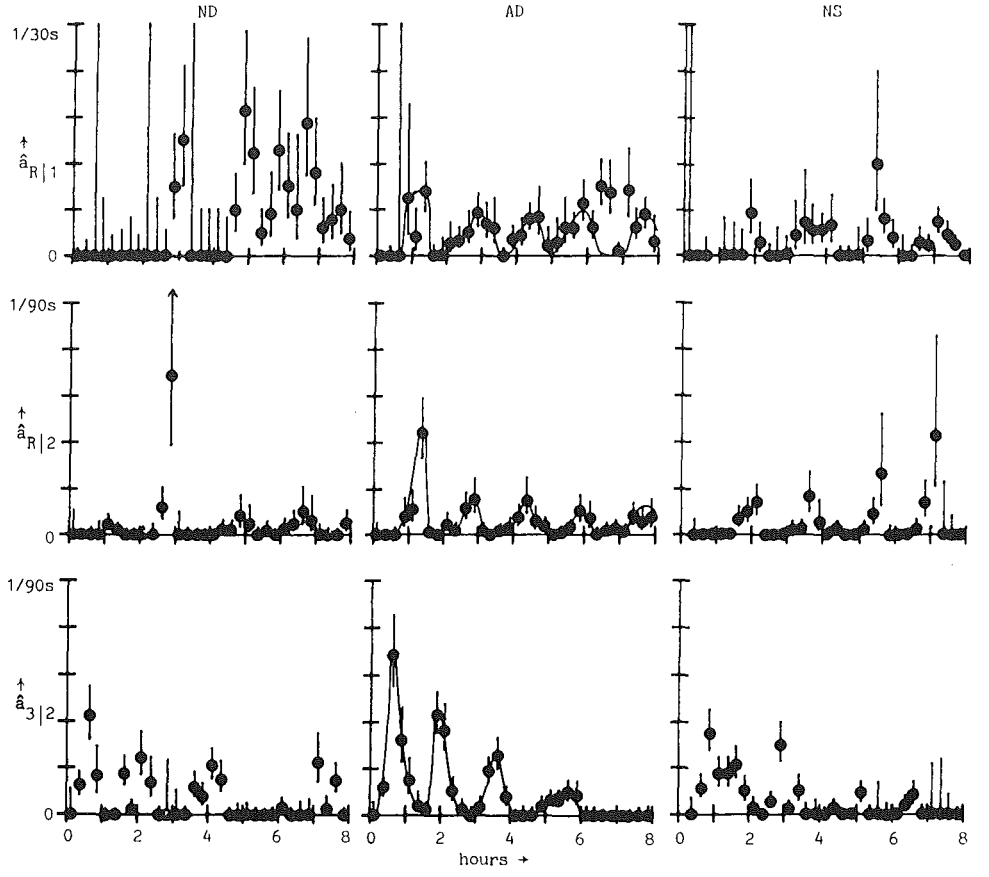


Figure 4: Individual time courses of transition rate estimates for subjects ND, AD and NS. Solid line, time course (by hand, AD only). Note the constant, synchronized periodicities in AD and NS with periods of ~ 93 and 101 minutes, respectively.

II.4. Simulations

In this section we describe how to generate simulations by the model. These enable both a general qualitative impression of its performance and statistical testing of any particular aspect. Some examples are given.

When sleep resides in stage i, there are six possible transitions to the other six stages, $j \neq i$. Each transition may occur with a certain probability, as specified by the transition rates, $a_{j|i}$. This corresponds to the transitions being generated by Poisson point processes (Snyder 1975:ch.2) with the same rates. As soon as one of these six Poisson processes generates a point, the corresponding transition, e.g. to stage $k \neq i$, occurs. Sleep now resides in stage k and the process starts anew.

The simulation based on Poisson point processes runs as follows. We start at time t at stage i. Six Poisson processes (with rates $a_{j|i}$) are active and they are simulated simultaneously as follows. A random number generator is used to obtain six independent variables, x_j, that are uniformly distributed in the interval $[0,1]$. A logarithmic transformation yields the variables, $-\ln(x_j)$, that are exponentially distributed. These would equal the six waiting times, $W_{j|i}$, for the first transitions to the six possible stages, $j \neq i$, if the Poisson processes were homogeneous with rates $a_{j|i}=1$. The correct waiting times for our inhomogeneous processes are obtained by a transformation of the time axis. These are the times, $W_{j|i}$, for which (Snyder 1975:ch.2)

$$\int_{t}^{t+W_{j|i}} a_{j|i}(\tau) \cdot d\tau = -\ln(x_j) \qquad (7)$$

After the shortest, $W_{k|i}$, of these six waiting times, a transition at time $t+W_{k|i}$ to stage k occurs. Thereafter, the simulation of the next transition starts.

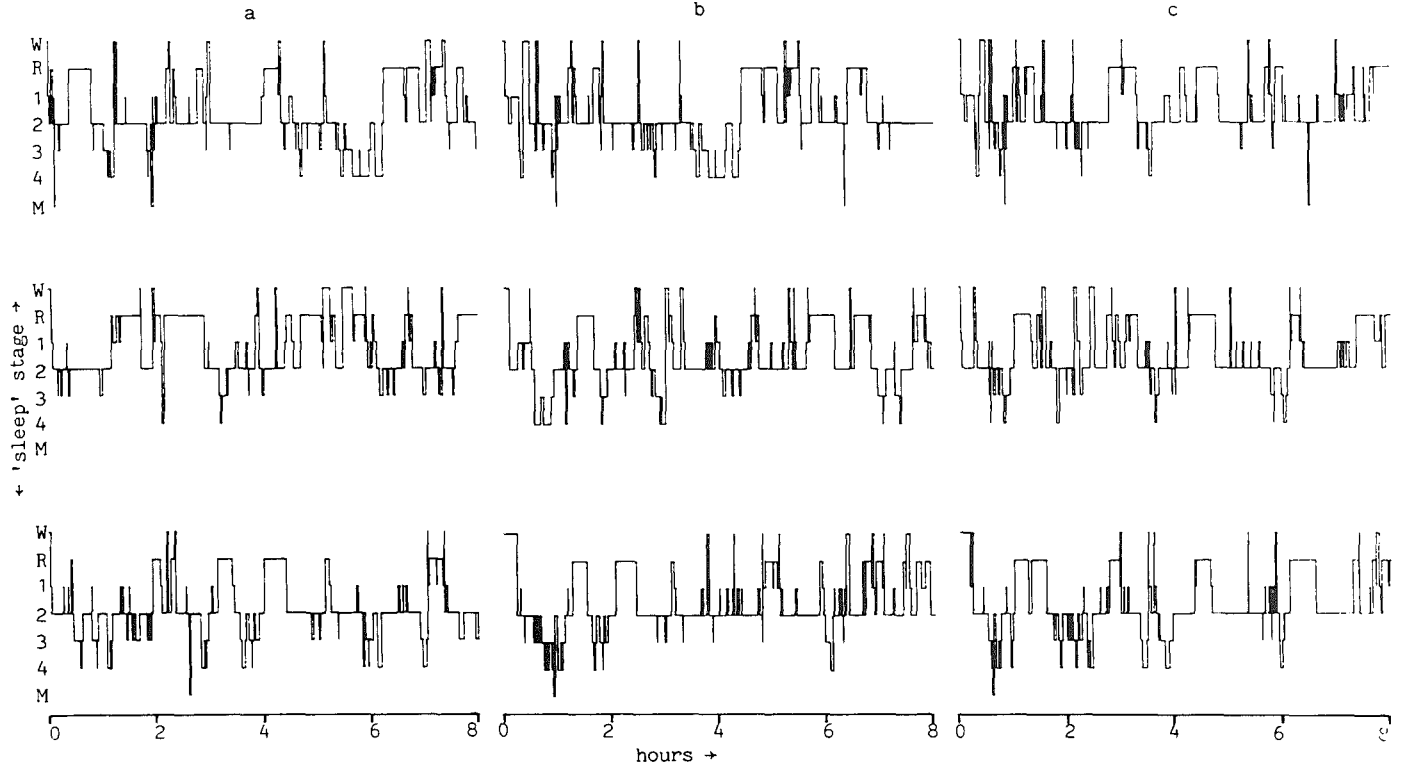


Figure 5: Simulated hypnograms, based on (a) constant whole group rates of table 1; (b) as a, but with the five smoothed time-varying whole-group rates of figure 3; (c) as b, but with the three individual smoothed rates of figure 4, subject AD. Note the absent, weak and comparable (to the real hypnograms of figure 1) REM-NREM periodicity in a, b and c, respectively.

Figure 5 shows three types of simulation. The first two merely serve as a further illustration, besides figures 3 and 4, to the experienced hypnogram-reader (others may use figure 1 as a reference) of the importance of time-variability and individuality in the model. Simulations that neglect these (fig.5^a and 5^b, respectively) can be seen, even at first glance, to differ clearly from real hypnograms. Only the 'individual' simulations in fig.5^c show a nightly trend in the amount of slow-wave-sleep and a REM-NREM periodicity comparable with real sleep.

II.5. Discussion

The proposed model suggests transition rates, $a_{j|i}$, as parameters that characterize the sleep mechanism. The rates reflect the well known trends and periodicities in sleep, as well as some interindividual differences. Simulations resemble real hypnograms. The arguments (Yang and Hursch 1973) against the model have been disproved. Therefore, the Markov model deserves reconsideration and this paper provides the methods for its specification, application and simulation.

The transition rates in our simulations can have any time course and they are set a priori and independent of each other, thus offering a very large degree of freedom. In fact, if the estimation interval, Δ, becomes infinitely small, the model will exactly reproduce the one hypnogram to which it has been adapted. Validation of the Markov, and of any more complex, model is therefore difficult. Although it may be possible to apply statistics individually, based on many nights per subject, this would still account for only some aspects. Therefore, simulated hypnograms should also be subjected to the more comprehensive judgment of experienced sleep researchers.

The model can probably be simplified because the rates seem to be generated by mechanisms that impose simple time courses (e.g. $a_{2|1}$, $a_{1|W}$ and

$a_{3|2}$), interdependencies between rates (e.g. $a_{R|2}$, $a_{R|1}$ and $a_{3|2}$ seem to be synchronized to each other), or sleep dependencies. Such mechanisms are also suggested by the deterministic sleep models mentioned in the introduction. It is possible to combine the Markov model and some deterministic models by interpreting the deterministic models as mechanisms that generate the rates for the Markov model. The result might be a model with few parameters that still simulates most probabilistic and deterministic aspects of sleep.

Acknowledgement: We hereby express our gratitude to Dr. A.N. Nicholson of the R.A.F. Institute of Aviation Medicine (U.K.), who kindly supplied us with the required high-quality sleep scores. We also gratefully acknowledge the constructive criticism and help of Professor Ir. E.W. Gröneveld of the Twente University of Technology and Professor Dr. F.H. Lopes Da Silva of the University of Amsterdam. Finally, we thank Ms. M.E. Van Hoogevest for the translation into French.

Résumé

Un processus de Markov a été proposé comme modèle pour le méchanisme qui produit les phases de sommeil humaines. Une méthode a été introduite pour l'estimation des paramètres du modèle, i.e. les probabilités de la transition (taux) entres les phases du sommeil, et a été appliquée aux 95 hypnogrammes chez 23 sujets. Les taux caractérisent des differences interindividuelles et des variations nocturnes du mécanisme du sommeil, lié au nombre diminuant du 'slow wave sleep' (vague lente de sommeil) dans le cours de la nuit, et la périodicité REM-NREM. Le modèle simule les deux, le dynamisme probable, ainsi que le dynamisme prédit sus-mentionné, mais seulement si ces taux individuels, variant en temps, sont appliqués.

III. AN OPTIMAL MONITOR OF THE ELECTROENCEPHALOGRAPHIC SIGMA SLEEP STATE

B. Kemp, P. Jaspers, J.M. Franzen, A.J.M.W. Janssen.

Reproduced from Biological Cybernetics 1985, 51:263-270

Summary

A model has been proposed for a Markov-jumping sleep depth that modulates a white-noise driven structure generating the sigma rhythm in the electroencephalogram. The corresponding maximum likelihood monitor, that continuously detects the current sleep stage from the observed electroencephalogram, has been derived and implemented. Simulations show high detection performances.

III.1. Introduction

The various stages of human sleep can be recognized by (a.o.) different eye or body movements, muscle tension and several components of the electroencephalogram (EEG)(Rechtschaffen and Kales 1968). One of the latter is the so-called sigma rhythm, a narrowband EEG component between 12 and 16Hz at C3, Cz, and C4 electrode locations (Jasper 1958), which is specific for light sleep (stage 2) and absent during wakefulness, drowsy sleep (stage 1) and REM sleep. Most systems for automatic sleep stage scoring are, therefore, partly based on sigma state monitoring.

A variety of sigma state monitors, based on several combinations of different techniques like phase-locked loops, bandpass filters, squarers, rectifiers, integrators, counters, percentage of time devices, Fourier transforms, complex demodulators, etc. (Pivic et al 1982, Larsen and Walter 1970, Gaillard and Blois 1981, Smith et al 1975, Broughton et al 1978, Johnson 1977, Vo-Ngoc et al 1971, Courtney and Noton 1972, Sciarretta and Bricolo

1970, Campbell et al 1980) is currently available. All these devices suggest their own definition of the sigma state; a definition with a technical rather than a physiological background. It is, therefore, difficult to decide which monitor is best (or sufficiently close) related to the 'real' physiological sigma state of the brain (Kemp 1981). System performances can only be evaluated by comparing them to human scoring and, therefore, may be suboptimal if the human decisions also are. Because the performance is likely to depend on the sigma power and on the average sleep stage duration, and both may vary considerably between different subjects, performance analyses are of limited value and very laborious.

The problem: 'how to extract the sigma sleep state from the recorded EEG', can also be approached using a precise, physiologically meaningful model of the sigma state parameter that has to be monitored, its relation to the generated EEG and a goodness of performance criterion. Derivation of the likelihood ratio for the thus defined problem then provides the optimal monitor for this parameter. Apparently this approach replaces the problem of finding the right monitor by the problem of finding the right problem formulation (a.o. EEG generator model), and therefore provides alternative selection criteria. Furthermore, the performance of the monitor may be established analytically or by simulations with the model. Also, reduced performance due to incompatibilities with the model, e.g. due to muscle or electrode artifacts, can be estimated by simulations with a mismatched model.

This paper describes the development of a sigma state monitor by this approach, its performance characteristics and some general conclusions with respect to sleep stage scoring.

III.2. Problem formulation: a model

We have conformed to the common proposition that a limited number of sleep stages exists and that during sleep discrete transitions between these stages occur. In the present paper we are only interested in transitions between sleep stages with and sleep stages without the electroencephalographic sigma rhythm.

A suitable description of these transitions and their relation to the EEG can be borrowed from a recently developed model of the alpha rhythm (Kemp and Blom 1981). This model (fig.1[a]) is based on the assumptions that alpha rhythms are oscillations which are caused by frequency dependent feedback loops, G(f), in noise driven cell networks. The feedback can be modulated by external influences, p(t), like visual attention or sleep. The transfer of cortical EEG to scalp EEG by volume conduction attenuates higher frequencies more than lower ones. This is accounted for by the lowpass filter, L(f). Simulations (fig.1[b]) show a waxing and waning (spindling) component when the feedback is active [i.e. p(t)=1]. This component vanishes by feedback inhibition [i.e. p(t)=0].

These rather general physiological assumptions have been proposed for the alpha rhythm (Lopes da Silva et al 1976), but may well also apply to the sigma rhythm (Shouse and Sterman 1979, Sterman and Bowersox 1981). The model accounts for the spindling character of the sigma rhythm. We have, therefore, adopted figure 1[a] as our EEG generator model, and the feedback modulator, $p(t) \in \{0,1\}$, as the sigma state that has to be monitored.

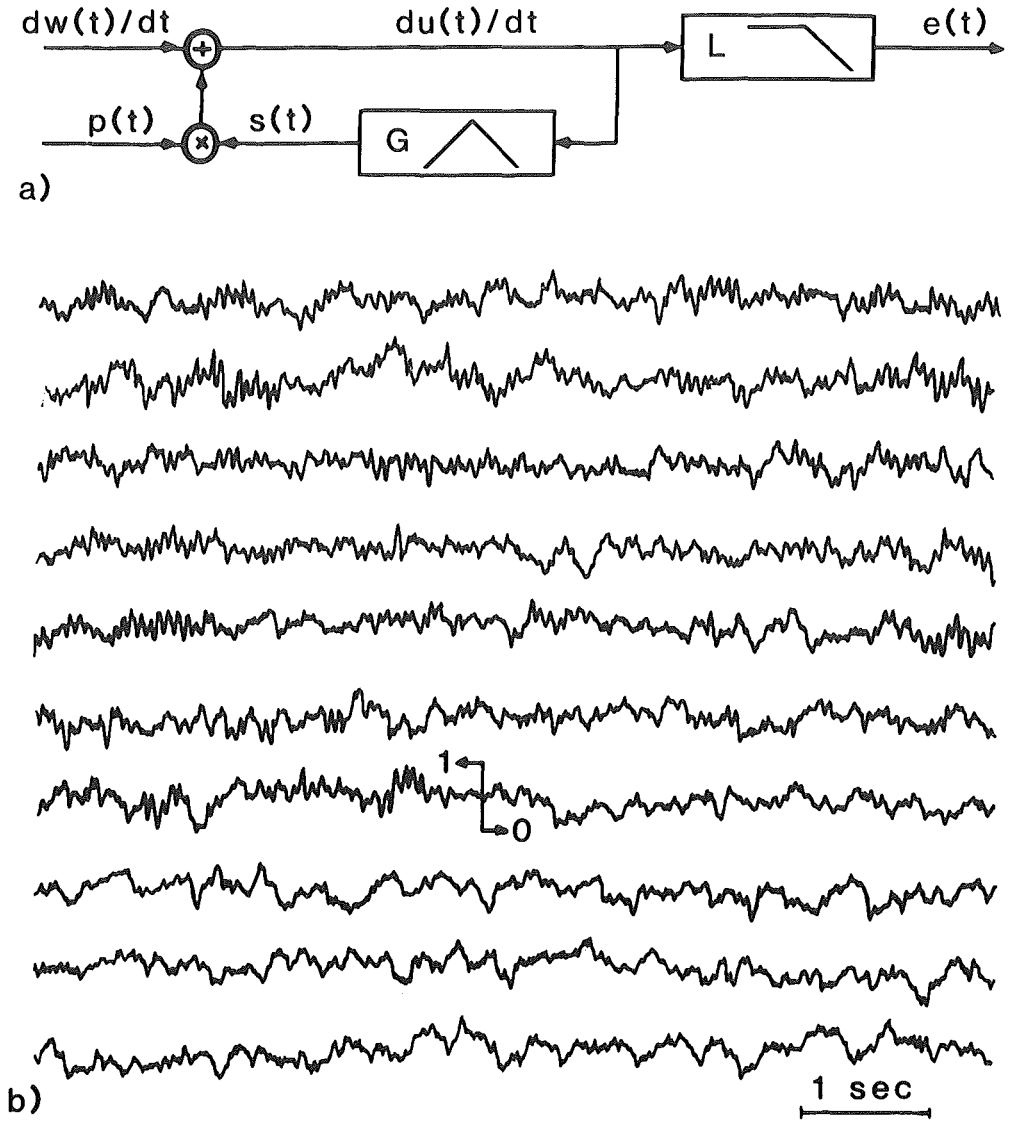


Figure 1: **a** Block diagram of EEG, e(t), generator model. Gaussian white noise, $d\omega(t)/dt$, and Markov feedback modulator, p(t). Second-order resonance filter, $G(f)=G_0/[1+j\cdot(f_0/B)\cdot(f/f_0-f_0/f)]$. First-order lowpass filter, $L(f)=L_0/[1+j\cdot(f/f_c)]$. **b** Simulation of 'EEG', e(t), before and after a feedback on-off [i.e. p(t)=1 to p(t)=0] transition. The transition occurs at the vertical bar. Parameter settings: $G_0=0.8$, B=4Hz, $f_0=14$Hz, $f_c=1.8$Hz.

Very little is known about the statistical properties of this sigma state. However, the sleep stage scoring guides which are used in clinical practice (e.g. Rechtschaffen and Kales 1968) do not invalidate the assumptions that the duration of sigma-on and sigma-off stages is independent of durations of preceding stages, and that the probability of a stage transition is not influenced by the duration of the present stage. If we take these two assumptions to be true, the transition probability densities (sigma state on-off and off-on transitions) are constant and as a consequence the sigma-on and sigma-off interval durations (T_1 and T_0, respectively) are mutually independent and exponentially distributed:

$$P\{T_1 \leq t\} = 1-\exp(-t/\tau_1) \quad (1)$$

$$P\{T_0 \leq t\} = 1-\exp(-t/\tau_0) \quad (2)$$

where $P\{.\}$ denotes a probability. The average $p(t)=1$ (sigma-on) and $p(t)=0$ (sigma-off) interval durations are τ_1 and τ_0, respectively. The transition probability densities mentioned above are $1/\tau_1$ and $1/\tau_0$, respectively.

These distributions have already been proposed by Zung et al (1965). They approximate the distributions of real sleep stage durations (Williams et al 1964, Yang and Hursch 1973, Brezinová 1975). Some examples are given in figure 2. Although the approximation errors may be significant (Yang and Hursch 1973), they are probably small compared to interindividual differences, the influence of sleep pathology or of the subjective scoring procedures (e.g. fig.2). They may also be caused by group inhomogeneities, since the average of different exponential distributions is not exponential. For these reasons, and because it is attractive from a mathematical point of view, we have adopted these exponential distributions as our sigma state generator model.

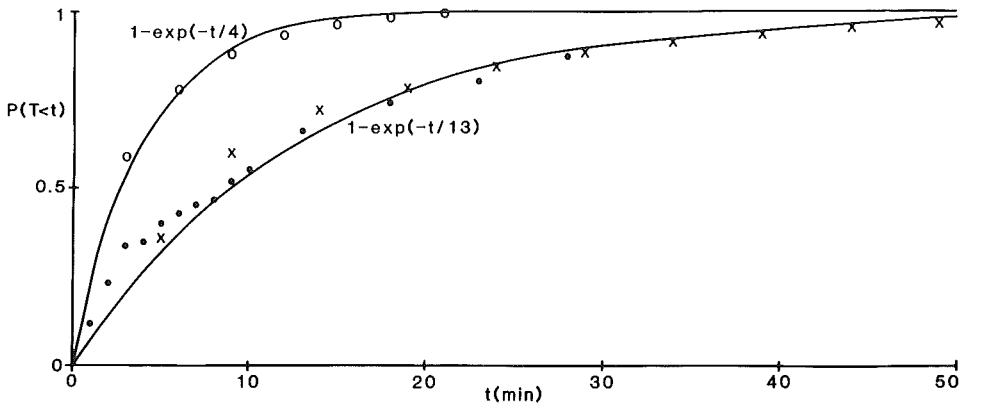


Figure 2: Distributions of durations, T, of sleep stage 2 intervals in healthy young subjects, as derived from original data of Yang and Hursch (•), Williams et al (x) and Březinová (o). The two fitted exponential distributions (solid lines) correspond to average interval durations of 13 and 4 min. The (•) and (x) groups have been scored by the same laboratory.

Based on this description of the sigma state, p(t), and its relation to the EEG, e(t), the problem can be reformulated as follows: which monitor provides at all times, t, during the recording interval, $\{t, 0 \leq t < T\}$, the optimal decision, $\overset{*}{p}(t)$, on the current value of p(t), based on the full EEG realisation, $E(0,T) = \{e(\nu), 0 \leq \nu < T\}$. As an optimization rule, we have adopted a Bayes criterion (Van Trees 1979): e.g. the monitor should minimize the expected false-decision rate.

III.3. Derivation of the sleep state monitor

First, we will show that the model of figure 1^a is equivalent to a set of two differential equations: one describing the process $p(\nu)$ as a differential equation driven by a martingale process, and the other describing its relationship to white noise observations. This description enables the use of a known filter algorithm (Wonham 1965) for the derivation of the monitor. The process $u(\nu)$ of figure 1^a can be reconstructed by inverse filtering by $L^{-1}(f)$

and integration of the recorded EEG, $e(\nu)$. From now on, $u(\nu)$ is taken to be the observed process. Because the feedback signal, $s(\nu)$, can be reconstructed by filtering $e(\nu)$ by $G(f)\cdot L^{-1}(f)$, it is a known signal. Therefore, figure 1^a is equivalent to the following differential equation, that relates the process $p(\nu)$ to the observations:

$$du(\nu) = p(\nu)\cdot s(\nu)\cdot d\nu + d\omega(\nu) \qquad (3)$$
$$u(0) = 0$$

where $u(\nu)$ are the observations, $s(\nu)$ is a known signal, $\omega(\nu)$ is a standard Wiener process [i.e.$d\omega(\nu)/d\nu$ is white noise] and $p(\nu)$ is the sigma state to be monitored. The process $p(\nu)$ can be written [in correspondence to (1) and (2)] as a binary Markov chain:

$$dp(\nu) = \{1-p(\nu)\}\cdot dq_0(\nu) - p(\nu)\cdot dq_1(\nu) \qquad (4)$$
$$p(0) \in \{0,1\}$$

where $q_1(\nu)$ and $q_0(\nu)$ are mutually independent Poison counting processes (Larson and Shubert 1979, Snyder 1975). These processes are staircase functions that are constant [i.e. $dq_i(\nu)=0$] except for positive unit steps [i.e. $dq_i(\nu)=1$] that occur with constant probability densities (also called rates), $1/\tau_1$ and $1/\tau_0$, respectively. Equation (4) can be transformed into:

$$dp(\nu) = \{(1-p(\nu))/\tau_0 - p(\nu)/\tau_1\}\cdot d\nu + dm(\nu) \qquad (5)$$
$$p(0) \in \{0,1\}$$

by substitution of:

$$dm(\nu) = \{1-p(\nu)\}\cdot\{dq_0(\nu)-d\nu/\tau_0\} - p(\nu)\cdot\{dq_1(\nu)-d\nu/\tau_1\} \qquad (6)$$

Because $\{dq_0(\nu)-d\nu/\tau_0\}$ and $\{dq_1(\nu)-d\nu/\tau_1\}$ are the increments of two centred Poisson processes (Snyder 1975), the process $m(\nu)$ is a martingale and (5) is a martingale-driven differential equation. A nice introduction to martingales has been given by Wong (1973). Any stochastic process, $m(\nu)$, with conditional expectation $\mathbf{E}\{m(\nu)|m(\tau),\tau\leq s\}=m(s)$ for all $s<\nu$, is a martingale. In (6), the expectations of the increments, $dq_i(\nu)$, are precisely compensated by the terms, $d\nu/\tau_i$. Therefore $\mathbf{E}\{dm(\nu)\}=0$ and $m(\nu)$ is a martingale.

Having described the model by the differential equations (3) and (5), we can now proceed to the derivation of the optimal monitor. The optimal decision, $\overset{*}{p}(t)$, on $p(t)$ can be obtained from the set of observations, $U(0,T)=\{du(\nu),0\leq\nu<T\}$, by comparing the likelihood ratio (Van Trees 1979):

$$L(t) = f\{U(0,T)|p(t)=1\}/f\{U(0,T)|p(t)=0\} \qquad (7)$$

where $f\{.|.\}$ are conditional probability densities, to a constant threshold, K. This threshold depends on the optimality criterion and will be discussed later. When $L(t)>K$ the optimal decision is 'sigma state on', i.e. $\overset{*}{p}(t)=1$. When $L(t)<K$ it is 'sigma state off', i.e. $\overset{*}{p}(t)=0$. Later, we will use the simple relationship (8) between $L(t)$ and the conditional expectation, $\hat{p}(t)=\mathbf{E}\{p(t)|U(0,T)\}$, wich can be derived by Bayes' theorem from (7):

$$\begin{aligned} L(t) &= [f\{U(0,T),p(t)=1\}/P\{p(t)=1\}] \cdot [P\{p(t)=0\}/f\{U(0,T),p(t)=0\}] = \\ &= [P\{p(t)=0\}/P\{p(t)=1\}] \cdot \\ &\cdot [P\{p(t)=1|U(0,T)\}/f\{U(0,T)\}] \cdot [f\{U(0,T)\}/P\{p(t)=0|U(0,T)\}] = \\ &= (\tau_0/\tau_1)\cdot\hat{p}(t)/\{1-\hat{p}(t)\} \end{aligned} \qquad (8)$$

where $P\{.\}$ denotes a probability, $P\{.|.\}$ a conditional probability and $f\{.\}$ a probability density. The likelihood ratio can be factorized into a 'past' observations part and a 'future' observations part as follows:

$$L(t) = \frac{f\{U(t,T)|U(0,t),p(t)=1\}\cdot f\{U(0,t)|p(t)=1\}}{f\{U(t,T)|U(0,t),p(t)=0\}\cdot f\{U(0,t)|p(t)=0\}} =$$

because of the Markov property of $p(\nu)$ and the independence of the Wiener increments, $d\omega(\nu)$, in (3)

$$= \frac{f\{U(t,T)|p(t)=1\}\cdot f\{U(0,t)|p(t)=1\}}{f\{U(t,T)|p(t)=0\}\cdot f\{U(0,t)|p(t)=0\}} =$$

$$= L^{+}(t)\cdot L^{-}(t) \qquad (9)$$

where $L^{+}(t)$ is the likelihood ratio based on only 'future' observations, $U(t,T)=\{du(\nu),t\leq\nu<T\}$ and where $L^{-}(t)$ is the likelihood ratio based on only 'past' observations, $U(0,t)=\{du(\nu),0\leq\nu<t\}$.

Like in Kemp (1983[a]) we will first concentrate on $L^{-}(t)$ which, like $L(t)$ from (7) and (8), can be written as a function of the conditional expectation, $\hat{p}^{-}(t)=\mathbf{E}\{p(t)|U(0,t)\}$:

$$L^{-}(t) = (\tau_0/\tau_1)\cdot\hat{p}^{-}(t)/\{1-\hat{p}^{-}(t)\} \qquad (10)$$

A differential equation to obtain the filtered estimate, $\hat{p}^{-}(t)$, has been derived by Wonham (1965) with the condition that $s(\nu)$ is non-anticipative with respect to $u(\nu)$ (see also Liptser and Shiryayev 1977). In order to fulfill this condition, the increments, $d\omega(\nu)$, in (3) should be interpreted as Itô-(forward-)differentials (see also Larson and Shubert 1979: ch.6). This interpretation does not affect the dynamics of the model (Wong and Zakai 1965). The filter differential equation reads:

$$d\hat{p}^{-}(\nu) = \{[1-\hat{p}^{-}(\nu)]/\tau_0 - \hat{p}^{-}(\nu)/\tau_1\}\cdot d\nu + \\ + s(\nu)\cdot\hat{p}^{-}(\nu)\cdot\{1-\hat{p}^{-}(\nu)\}\cdot\{du(\nu)-\hat{p}^{-}(\nu)\cdot s(\nu)\cdot d\nu\} \qquad (11)$$

$$\hat{p}^{-}(0) = \tau_1/(\tau_0+\tau_1)$$

Driving this equation from $\nu=0$ to $\nu=t$ yields $\hat{p}^{-}(t)$. Equations (11) and (10) are theoretically sufficient for the recursive computation of $L^{-}(t)$. However, the proper physical implementation of (11) is obscured (Wonham 1965, Kailath 1969, Clements and Anderson 1973) because of the unknown behaviour of the white-noise-containing last term, $s(\nu)\cdot\hat{p}^{-}(\nu)\cdot\{1-\hat{p}^{-}(\nu)\}\cdot\{du(\nu)-\hat{p}^{-}(\nu)\cdot s(\nu)\cdot d\nu\}$. The effects of discretizing the algorithm, limiting the frequency band and high frequency deviations from the model are not quite clear. This difficulty can be met by a logarithmic transformation of L(t). For digital implementation reasons (sect.III.4) we have chosen a $^2\log$ transformation, yielding the new variable, $\Lambda(t)$, also called the test statistic for reasons to be explained later:

$$\Lambda(t) = {}^2\log\{L(t)\} = {}^2\log\{L^{-}(t)\} + {}^2\log\{L^{+}(t)\} \\ = \Lambda^{-}(t) + \Lambda^{+}(t) \qquad (12)$$

where

$$\Lambda^{-}(t) = {}^2\log\{L^{-}(t)\} = {}^2\log\{(\tau_0/\tau_1)\cdot\hat{p}^{-}(t)/[1-\hat{p}^{-}(t)]\} \qquad (13)$$

The differential equation for the computation of $\Lambda^{-}(t)$ can be derived from (13) and (11). Because (11) is driven by white noise, this may not be done by the ordinary chain rule for differentiation, but Itô's differentiation rule (Larson and Shubert 1979, Kailath 1969) should be applied as follows:

$$d\Lambda^{-}(\nu) = [\delta\Lambda^{-}(\nu)/\delta\hat{p}^{-}(\nu)]\cdot d\hat{p}^{-}(\nu) +$$

$$+ \tfrac{1}{2}[\delta^{2}\Lambda^{-}(\nu)/\delta\hat{p}^{-2}(\nu)]\cdot\{s(\nu)\cdot\hat{p}^{-}(\nu)\cdot[1-\hat{p}^{-}(\nu)]\}^{2}\cdot d\nu \qquad (14)$$

The partial derivatives can be obtained from (13):

$$\delta\Lambda^{-}(\nu)/\delta\hat{p}^{-}(\nu) = {}^{2}\log(e)/\{\hat{p}^{-}(\nu)\cdot[1-\hat{p}^{-}(\nu)]\} \qquad (14^{a})$$

$$\delta^{2}\Lambda^{-}(\nu)/\delta\hat{p}^{-2}(\nu) = {}^{2}\log(e)\cdot\{2\hat{p}^{-}(\nu)-1\}/\{\hat{p}^{-}(\nu)\cdot[1-\hat{p}^{-}(\nu)]\}^{2} \qquad (14^{b})$$

Substitution in (14) of these derivatives and of $d\hat{p}^{-}(\nu)$ from (11) yields the differential equation for $\Lambda^{-}(t)$:

$$d\Lambda^{-}(\nu) = {}^{2}\log(e)\cdot[[1/\{\tau_{0}\cdot\hat{p}^{-}(\nu)\}-1/\{\tau_{1}\cdot[1-\hat{p}^{-}(\nu)]\}]\cdot d\nu +$$

$$+ s(\nu)\cdot\{du(\nu)-\hat{p}^{-}(\nu)\cdot s(\nu)\cdot d\nu\} + \tfrac{1}{2}s^{2}(\nu)\cdot\{2\hat{p}^{-}(\nu)-1\}]\cdot d\nu =$$

$$= {}^{2}\log(e)\cdot[[\{2^{-\Lambda^{-}(\nu)}-1\}/\tau_{1}-\{2^{\Lambda^{-}(\nu)}-1\}/\tau_{0}]d\nu + s(\nu)du(\nu)-\tfrac{1}{2}s^{2}(\nu)d\nu]$$

$$\Lambda^{-}(0) = 0 \qquad (15)$$

Driving this equation from $\nu=0$ to $\nu=t$ yields $\Lambda^{-}(t)$. Figure 3 shows the implementation block diagram. It appears to be an integrator with exponential feedback. The interpretation of the white noise containing driving term, $dc(\nu)={}^{2}\log(e)\{s(\nu)du(\nu)-\tfrac{1}{2}s^{2}(\nu)d\nu\}$, is now straightforward. From (3), we have:

$$dc(\nu) = {}^{2}\log(e)\cdot\{s(\nu)\cdot d\omega(\nu)+[p(\nu)-\tfrac{1}{2}]\cdot s^{2}(\nu)\cdot d\nu\}$$

Because $d\omega(\nu)$ is an independent forward increment, the conditional expectation of the driving term equals:

$$E\{dc(\nu)|p(\nu)\} = {}^2\log(e)\cdot\{[p(\nu)-\tfrac{1}{2}]\cdot s^2(\nu)\cdot d\nu\}$$

Therefore, the integrator will be driven to positive values during sigma-on intervals [i.e. $p(\nu)=1$] and to negative values during sigma-off intervals [i.e. $p(\nu)=0$]. The saturation effect, caused by the exponential feedback, limits the time needed to react to a transition of $p(\nu)$. Quite satisfactory, this effect increases with increasing transition rates, $1/\tau_1$ and $1/\tau_0$. An equivalent 'past' observations test has already been derived (Kemp and Blom 1981) by solving the discrete time analogue of the problem.

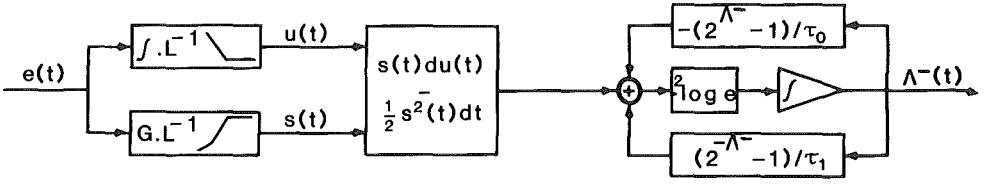


Figure 3: Implementation block diagram for the generation of the 'past' observations test statistic, $\Lambda^-(t)$.

Because (5) may also be interpreted as a backward differential equation, the transformed 'future' test statistic, $\Lambda^+(t)$, follows from exactly the same differential equation:

$$d\Lambda^+(\nu) = {}^2\log(e)\cdot[[\{2^{-\Lambda^+(\nu)}-1\}/\tau_1-\{2^{\Lambda^+(\nu)}-1\}/\tau_0]d\nu + s(\nu)du(\nu)-\tfrac{1}{2}s^2(\nu)d\nu]$$
$$\Lambda^+(T) = 0 \qquad (16)$$

except that this is a backward differential equation that has to be driven from $\nu=T$ ('future') to $\nu=t$ ('present') in order to obtain $\Lambda^+(t)$.

According to (9), summation of $\Lambda^{-}(t)$ and $\Lambda^{+}(t)$ yields the desired test statistic, $\Lambda(t)$, from which the likelihood ratio, $L(t)$, can be obtained. However, instead of comparing $L(t)$ to the threshold, K, in order to obtain the optimal decision, $\overset{*}{p}(t)$, a more direct procedure is to compare $\Lambda(t)=^{2}\log\{L(t)\}$ directly to the transformed threshold, $\xi=^{2}\log\{K\}$. In other words, $\Lambda(t)$ is a sufficient statistic for testing whether $p(t)=1$ or $p(t)=0$. Generally, the Bayes' threshold for minimizing expected 'costs' reads (Van Trees 1979):

$$K = (C_{1|0}/C_{0|1})\cdot P\{p(t)=0\}/P\{p(t)1\} = \\ = (C_{1|0}/C_{0|1})\cdot(\tau_0/\tau_1) \qquad (17)$$

where $C_{1|0}$ and $C_{0|1}$ are the 'costs' of the false decisions, $\overset{*}{p}(t)=1$ and $\overset{*}{p}(t)=0$, respectively. If we want to minimize the false decision rate (i.e. $C_{1|0}=C_{0|1}$), the transformed threshold equals:

$$\xi = {}^{2}\log\{K_{rate}\} = {}^{2}\log\{\tau_0/\tau_1\} \qquad (18)$$

If we want to minimize the false decision percentages (i.e. $C_{1|0}/C_{0|1}=\tau_1/\tau_0$), this threshold equals:

$$\xi = {}^{2}\log\{K_{perc}\} = 0 \qquad (19)$$

III.4. Implementation

We will not describe our implementation in full detail. However, a few remarks can be made that will facilitate any particular application.

Because of the monitor's non-linear feedback (fig.3), the signal power is important. The reconstructed signals should be amplified so that, during sigma-off intervals (e.g. during drowsy sleep or wakefulness), the

reconstructed observations, $u(\nu)$, equal a standard Wiener process, i.e. its derivative is white noise with two-sided power spectral density 1.

The resonance filter, G(f), which is required for the reconstruction of $s(\nu)$, is basically an LRC-circuit (fig.4) in which the inductance, L, has been realized by gyrating a capacitance, C_1. The gyrator consists of two matched operational transconductance amplifiers (LM13700). Both transconductances, g, are linearly proportional to the bias current, i_g. Therefore, the resonance frequency, $f_0=1/(2\pi\sqrt{LC})=g/(2\pi C_1)$, can be tuned by a linear potentiometer with a linear dial. The bandwidth, $B=(1/R_1+1/R_2)/(2\pi C_1)$, has been fixed at 4Hz. The maximum gain, G_0, can be tuned linearly from 0 to 1 by the potentiometer, R.

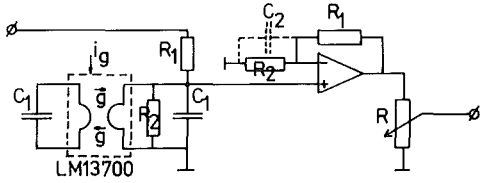


Figure 4: Analogue realization of the resonance filter, G(f). By the capacitor, C_2, the circuit also performs the inverse filtering by $L^{-1}(f)$.

The inverse filtering by $L^{-1}(f)$ in the signal path for reconstruction of $s(\nu)$ may be performed by analogue circuitry (e.g. by adding capacitor C_2 in figure 4). The reconstruction of $du(\nu)$, however, must be approximated in discrete time as follows:

$$\Delta u(\nu) = e(\nu).\Delta + \{e(\nu+\Delta)-e(\nu)\}/2\pi f_c \qquad (20)$$

Using this implementation, it is not necessary to reconstruct the physically not realizable white noise containing process $du(\nu)/d\nu$. Also, by the forward time shift, Δ, the differences, $\Delta u(\nu)$, are forward differences with respect

to $s(\nu)$. They approximate the differentials, $du(\nu)$, that were required to be forward differentials in the derivation of (11). Only with this forward time shift, does the discrete time approximation, $\Delta c(\nu)={}^2\log(e)\{s(\nu)\Delta u(\nu)-\frac{1}{2}s^2(\nu)\Delta\}$, behave like the continuous time driving term, $dc(\nu)$, in the original equations (15) and (16), i.e.:

$$\mathbf{E}\{\Delta c(\nu)|p(\nu)\} = {}^2\log(e)\cdot\{[p(\nu)-\tfrac{1}{2}]\cdot s^2(\nu)\cdot\Delta\}$$

Not implementing this shift will cause severe bias (Kemp 1983[b]).

By choosing a time step $\Delta={}^e\log(2)\cdot 2^{-6}$s, the discrete time approximations of (15) and (16) are the simple difference equations:

$$\Delta\Lambda^i(\nu) = [\{2^{-\Lambda^i(\nu)}-1\}/\tau_1-\{2^{\Lambda^i(\nu)}-1\}/\tau_0]\cdot 2^{-6} + \Delta c(\nu) \qquad (21)$$

Where $i\epsilon\{-,+\}$ and the driving term:

$$\Delta c(\nu) = s(\nu)\cdot[e(\nu)\cdot 2^{-6}+\{e(\nu+\Delta)-e(\nu)\}\cdot{}^2\log(e)/(2\pi f_c)-\tfrac{1}{2}s(\nu)\cdot 2^{-6}] \qquad (22)$$

Because the a priori informations, τ_0, τ_1, and f_c, are rather crude estimates, we may as well require τ_0, τ_1, and ${}^2\log(e)/(2\pi f_c)$ to be powers of 2. Now, one update, $\Delta\Lambda^i(\nu)$, requires only bit sets, bit shifts, adding and one multiplication.

The forward equation (21,i=-) can be updated on-line. The backward equation (21,i=+) requires storage of $\Delta c(\nu)$ over de 'future' interval, $\{t\leqslant\nu<T\}$. However, because of the exponential feedback, the length of this interval may be limited to 54s, i.e. $\{t\leqslant\nu<t+54\}$, without noticeable performance degrading.

All digital operations can be performed in real time by a MC6801-based microcomputer unit. The total parts costs of the (analogue and digital) system (fig.5) are about US$400.-

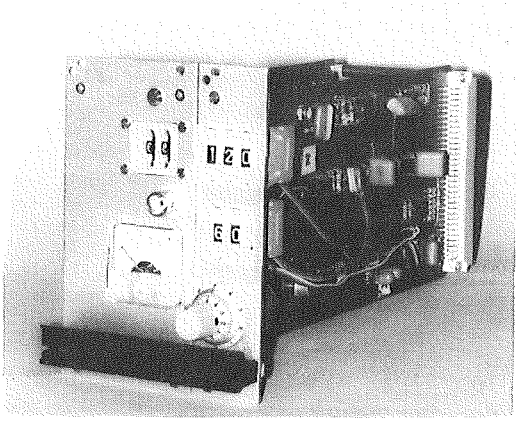


Figure 5: Realization of the monitor, showing continuous control for tuning the signal power and thumbwheels for setting G_0, f_0, τ_0 and τ_1.

III.5. Performance

A rather simple performance analysis (fig.6) can be carried out by simulating a patient using our EEG generator model. Although the obtained values will not be valid for real patients, some important qualitative conclusions can be drawn.

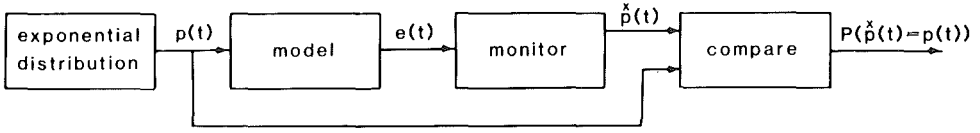


Figure 6: Performance analysis system

The modulating process, p(t), has been simulated according to the described exponential distributions (method in Snyder 1975:p.62) and fed to the analogue model of figure 1[a]. The output 'EEG', e(t), of the model is then fed into the sigma state monitor that provides the optimal decision, $\overset{*}{p}(t)$, on

$p(t)$. Comparing $\overset{*}{p}(t)$ to $p(t)$ over a period of 15h gives the probability of a correct decision, $P\{\overset{*}{p}(t)=p(t)\}$. For a 'difficult' patient' with a feedback gain, G_0=0.6, corresponding to very low sigma power, and with average interval lengths, $\tau_1=\tau_0$=64s, corresponding to rather many state transitions, this probability is 0.989. Most false decisions were made near interval boundaries. The results for other gains and interval lengths, including some in which the monitor parameters differ from the model parameters, have been summarized in table 1.

	Model				Monitor				Performance
	G_0	f_0	τ_0	τ_1	G_0	f_0	τ_0	τ_1	$P\{\overset{*}{p}(t)=p(t)\}$
X	0.60	12	64	64	0.60	12	64	64	0.989
Matched gains	0.70	12	64	64	0.70	12	64	64	0.990
	0.30	12	64	64	0.30	12	64	64	0.947
	0.20	12	64	64	0.20	12	64	64	0.891
	0.10	12	64	64	0.10	12	64	64	0.745
	0.05	12	64	64	0.05	12	64	64	0.608
Matched rates	0.60	12	1	1	0.60	12	1	1	0.747
	0.60	12	4	4	0.60	12	4	4	0.873
	0.60	12	16	16	0.60	12	16	16	0.958
	0.60	12	256	256	0.60	12	256	256	0.997
Unmatched gains	0.10	12	64	64	0.60	12	64	64	0.501
	0.30	12	64	64	0.60	12	64	64	0.770
	0.50	12	64	64	0.60	12	64	64	0.983
Unmatched rates	0.60	12	16	16	0.60	12	64	64	0.948
	0.60	12	16	64	0.60	12	64	64	0.980
	0.60	12	64	16	0.60	12	64	64	0.975
	0.60	12	256	256	0.60	12	64	64	0.997
Unmatched frequencies	0.60	8	64	64	0.60	12	64	64	0.502
	0.60	9	64	64	0.60	12	64	64	0.694
	0.60	10	64	64	0.60	12	64	64	0.973
	0.60	16	64	64	0.60	12	64	64	0.535

Table 1: Performance for different (matched and unmatched) values of gain, G_0, centre frequency, f_0(Hz), and rates, τ_1 and τ_0 (s), of the model and the detector. X values are standard settings, corresponding to a 'difficult' patient.

III.6. Discussion

Apparently it is possible to relate a model, that describes the most conspicuous characteristics of the EEG and the sleep stages, to the one and only optimal sleep stage monitor for that model. Therefore, optimizing the model implies optimization of the monitor. For instance, if non-linearities in the sigma rhythm generation can be simulated by a non-linearity in the bandpass filter, G(f), this non-linearity should also be incorporated into the monitor's G(f). Also, the exponential feedback of the integrator can be optimized by tuning τ_1 and τ_0 to the real, possibly time-dependent statistics of the sleep stages.

The driving correlator term, dc(ν), of the integrator suggests that power measures are more appropriate than amplitude (or envelope) measures. In fact, the sometimes applied bandfilter-squarers come rather close, although they generally have smaller bandwidths. The exponential feedback integrator provides an answer to the segmentation problem that arises when using the standard scoring rules (Rechtschaffen and Kales 1968) or methods like power spectral analysis. On the one hand, segments should be short in order not to miss small intervals, but on the other hand they should be long in order to minimize variance. The optimal compromise appears to be the exponential feedback integrator. The decision threshold, K, provides the possibility of obtaining different results for different problems like minimizing the false-decision rate vs. minimizing false-decision percentages.

For practical sleep stage scoring, some interesting conclusions can already be drawn. Firstly, mathematics may create original and attractive solutions to this old problem, when applied to suitable models. Secondly, the noisy character of the EEG and the jumping behaviour of the sleep stages do not necessary abolish the possibility of obtaining 99% correct decisions. And finally, monitoring performance is hardly degraded by the small mistunings of

sigma power and sleep stage statistics that will inevitably occur in practice.

Acknowledgements. We gratefully acknowledge the constructive criticism and help of Dr. J.H. Van Schuppen of the Centre for Mathematics and Computer Science in Amsterdam, Professor Dr. F.H. Lopes Da Silva of the University of Amsterdam, Professor Ir. E.W. Gröneveld of the Twente University of Technology, and Professor Dr. H.A.C. Kamphuisen of our own department.

IV. AN OPTIMAL MONITOR OF THE RAPID-EYE-MOVEMENT BRAIN STATE

B. Kemp

Reproduced from Biological Cybernetics 1986, 54:133-139

Summary

A model has been proposed for the stochastic occurrence of bursts of rapid eye movements (REMs) during sleep. REM-bursts are simulated by a Poisson counting process with a rate that depends on a binary Markov 'sleep state'. The corresponding maximum likelihood detector, that continuously monitors the current sleep state based on the observed REM-bursts, has been derived and implemented. Simulations show 97% correct decisions.

IV.1. Introduction

The various stages of human sleep can be recognized by (a.o.) different eye or body movements, muscle tension and several components of the electroencephalogram (Rechtschaffen and Kales 1968). One of the former are the rapid eye movements (REMs): fast rotations of the eyes which occur irregularly, but almost exclusively during wakefulness and during one of the sleep stages that is consequently called REM-sleep (Herman et al 1983, Schiller 1984). Most systems for automatic sleep stage monitoring are, therefore, partly based on monitoring the 'REM-state' of the brain; that is the state (either wakefulness or REM-sleep) during which REMs do occur.

Figure 1: Eye movement recording during REM-sleep. Note the two bursts of Rapid-Eye-Movements (+).

The eyes form strong dipoles. Their REMs occur abruptly and are much faster than other eye movements which may also occur during other (NREM) sleep stages (Herman et al 1983). Both eyes move simultaneously in the same direction. REMs can, therefore, be detected fairly reliably, either visually or automatically, from recordings using surface electrodes in the vicinity of the eyes (e.g. Broughton 1982, McPartland and Kupfer 1978, Ktonas and Smith 1978, Toth 1971, Goldberg and Beiber 1979, Gopal and Haddad 1981, Kayed et al 1979, Smith 1978 and figure 1). However, because the time between successive REMs during an uninterrupted REM brain state may amount to several minutes (Jacobs et al 1971, Aserinsky 1971, Dement and Kleitman 1957), it is difficult to decide whether there has been a real interruption of the REM-state by a NREM-state or whether REMs were accidentally not generated for some time during an uninterrupted REM-state. Some type of temporal smoothing must provide the appropriate decision. For instance, the usually applied standard sleep scoring rules state that any interval without REMs, but contiguous with stage REM, should be classified as stage REM except when clear signs for other stages do occur. It is not clear how many stage interruptions will thus be ignored (Rechtschaffen and Kales 1968).

Because both the brain state and the REMs during the REM state seem to be stochastic processes, an appropriate smoother should take into account their statistical properties. Such a smoother can be developed by modelling those processes and deriving the likelihood ratio for the problem 'state REM or state NREM?'. Using this approach we recently developed a brain state monitor based on continuous electroencephalographic (EEG) observations (Kemp et al 1985). It included an integrator with exponential feedback as a smoother that was based on brain state statistics. The integrator was driven by the EEG. However, here we are dealing with REMs that are discontinuous point observations. Therefore the same mathematics do not apply. It is not

clear whether this type of integrator may be used and if so, how it should be driven.

In this paper we propose a model for the REM observations and derive the corresponding maximum likelihood monitor for the underlying brain state. We describe its performance characteristics and some general implications for sleep stage classification.

IV.2. Problem formulation: a model

Both the brain state and the REM generating mechanism will be formulated in the form of differential equations that are driven by martingale processes. This description enables the use of a known filter algorithm for the derivation of the monitor.

We have conformed to the usual classification of sleep into a limited number of stages. Little is known about the statistical properties of the transitions between them. However, the sleep stage scoring guides that are used in clinical practice (e.g. Rechtschaffen and Kales 1968) do not invalidate the assumption that there is, at any moment, a fixed probability for the occurrence of each possible transition. If we take this assumption to be true, sleep stages are generated by a continuous time Markov chain process (e.g Larson and Shubert 1979). This process has already been proposed as a sleep stage generator model by Zung et al (1965) and its simulations show a good resemblance to real sleep stage patterns (Kemp and Kamphuisen 1986). If the process were homogeneous, i.e. with constant transition probabilities, then the sojourn times (being the times between successive transitions, during which sleep remains in the same stage) would be exponentially distributed. These distributions indeed approximate those of real sleep stage sojourn times (Williams et al 1964, Yang and Hursch 1973, Březinová 1975, Kemp et al 1985). For these reasons, and because it is attractive from a

mathematical point of view, we have adopted the Markov chain process as our brain state generator model.

In the present paper we are interested only in the transitions between REM states (many REMs) and NREM states (few or no REMs). We have therefore simplified the Markov chain to a binary one, $p(t)\epsilon(0,1)$, and REMs are generated predominantly when $p(t)=1$. The sleep state generating mechanism shows clear periodicities and trends that are different for different individuals (Kemp and Kamphuisen 1986). Since one generally does not dispose of sufficiently reliable a priori information about these individual dynamics, we have further simplified the binary Markov chain to one with constant transition rates, i.e. a homogeneous one. Another argument for this is that we intend to use the monitor for the quantification of these dynamics and this should not be biased by a priori assumptions. Later, we will show that this choice is not very critical for the performance of the monitor. The corresponding homogeneous brain state generating differential equation reads (Kemp et al 1985):

$$
\begin{aligned}
dp(\theta) &= [1-p(\theta)]\cdot dq_0(\theta) - p(\theta)\cdot dq_1(\theta) = \\
&= \{[1-p(\theta)]/\tau_0 - p(\theta)/\tau_1\}d\theta + [1-p(\theta)][dq_0(\theta) - d\theta/\tau_0] - p(\theta)[dq_1 - d\theta/\tau_1] = \\
&= \{[1-p(\theta)]/\tau_0 - p(\theta)/\tau_1\}d\theta + dm_1(\theta) \qquad (1) \\
p(0) &\in \{0,1\}
\end{aligned}
$$

where $q_1(\theta)$ and $q_0(\theta)$ are mutually independent Poisson counting processes. These processes are constant (i.e. $dq_i(\theta)=0$) except for positive unit counts (i.e. $dq_i(\theta)=1$) that occur with rates $1/\tau_1$ and $1/\tau_0$, respectively. Consequently, $m_1(\theta)$ according to (1) is a martingale. The average sojourn times in the REM state ($p(\theta)=1$) and the NREM state ($p(\theta)=0$) are τ_1 and τ_0 (about 20 min. and 60 min.), respectively.

Although there is some correlation between REMs and activity in certain neuroanatomically defined structures (Nelson et al 1983, McCarley et al 1983, Friedman and Jones 1984, Alihanka et al 1975), the physiological mechanism underlying their irregular appearance is not known. However, a phenomenological model may be based upon the statistics of the waiting times between two successive REMs. The REMs tend to occur in short bursts and several authors (Spreng et al 1968, Aserinsky 1971, Jacobs et al 1971) have demonstrated two groups of waiting times during the REM state. The first are intraburst times, shorter than about 3 seconds. The other are interburst times, averaging about 1 minute. Because these rather long times between bursts within the REM state are critical for REM state monitoring, they will form the basis for our model. The same authors suggest the possibility of an exponentially distributed waiting time, especially when intraburst times are omitted. They present waiting time histograms that support this suggestion. Although they show significant differences to this distribution, these may be caused by the intra- and interindividual variability in REM density that is known to exist (Feinberg et al 1967, Salzarulo 1972, Spiegel 1981). Therefore we have adopted the Poisson counting process, $n(\theta)$, that is suggested by these distributions, as a reasonable first approximation of the REM-burst statistics. Each count, $dn(\theta)=1$, generates a REM-burst and the intercount intervals are exponentially distributed. As in (1) we can write the REM-burst counting process in the form of a martingale-driven differential equation:

$$\begin{aligned} dn(\theta) &= r(\theta)\cdot d\theta + dn(\theta)-r(\theta)\cdot d\theta \\ &= r(\theta)\cdot d\theta + dm_2(\theta) \\ n(0) &= 0 \end{aligned} \qquad (2)$$

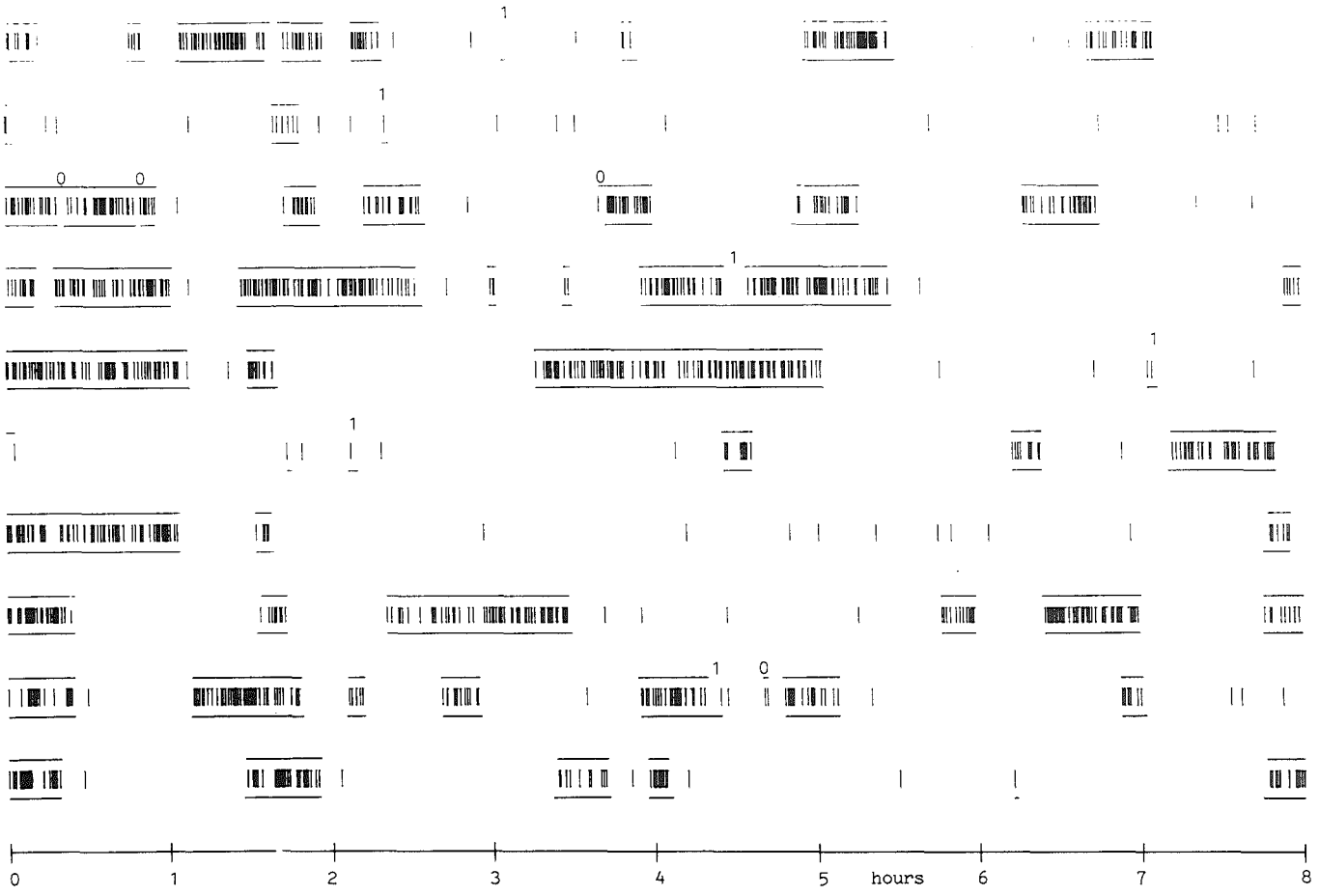
where $r(\theta)$ is the rate of the Poisson process, $n(\theta)$, i.e. the density of the

REM-bursts. Consequently, $m_2(\theta)$ is a martingale. During REM states ($p(\theta)=1$) this rate equals r_1 (about 1/min). During NREM states ($p(\theta)=0$), the rate is partly determined by false positive REM detections and equals r_0 (about .02/min). Or equivalently:

$$r(\theta) = r_0 + (r_1-r_0)\cdot p(\theta) \qquad (3)$$

Equation (1) describes the generation of the brain state, while (3) and (2) describe the related generation of the observed REM-bursts, $dn(\theta)$. Figure 2 shows a simulation by this model. Based on this description, the problem can be reformulated as follows: find the monitor which provides at all times, t, during the recording interval $\{t,0\leq t<T\}$, the optimal decision, $\overset{*}{p}(t)$, on the current value of p(t), based on the whole night observation, $N(0,T)=\{dn(\theta),0\leq\theta<T\}$, of REM bursts, $dn(\theta)$. As an optimization rule, we have adopted a Bayes criterion (e.g. Van Trees 1979): the monitor should minimize the expected false-decision rate.

Next page:
Figure 2: Monitoring the model during 10 'nights' of 8 hours. The upper, middle and lower traces of each night show:
—— ------ : optimal decision is 'REM', when interrupted 'NREM'
‖ ‖ ‖‖ : 9 simulated REM-bursts, dn(t)
—— —————— : simulated brain state is 'REM', when interrupted 'NREM'
Parameters of both the model and the monitor: τ_1=20 min., τ_0=60 min., r_1=1/min., r_0=.02/min. Note some incorrect classifications of the REM-state (1) and the NREM-state (0).


1
1
0
0
0
1
1
1
1
0
0
1
2
3
4
5
hours
6
7
8

IV.3. Derivation of the REM state monitor

The optimal decision, $\overset{*}{p}(t)$, on p(t) can be obtained from the set of observations, N(0,T), by comparing the likelihood ratio (Van Trees, 1979):

$$L(t) = f\{N(0,T)|p(t)=1\}/f\{N(0,T)|p(t)=0\} \qquad (4)$$

where $f\{.|.\}$ are conditional probability densities, to a constant threshold, K. K depends on the optimality criterion. We adopted the Bayes' threshold for minimizing the expected false-decision rate, which reads (Kemp et al 1985):

$$K=\tau_0/\tau_1 \qquad (5)$$

When L(t)>K the optimal decision is 'brain state REM', i.e. $\overset{*}{p}(t)=1$. When L(t)<K it is 'brain state NREM', i.e. $\overset{*}{p}(t)=0$. Using a similar reasoning as in our earlier paper (Kemp et al 1985), but now using the Markov property of $p(\theta)$ and the mutual independency of the Poisson counts, $dn(\theta)$, we may split L(t) into a 'future observations' part and a 'past observations' part as follows:

$$L(t) = \{L^-(t)\}\cdot\{L^+(t)\} \qquad (6)$$

and again:

$$L^i(t) = (\tau_0/\tau_1)\cdot\hat{p}^i(t)/[1-\hat{p}^i(t)] \qquad i\epsilon(-,+) \qquad (7)$$

where the expectations, based on 'past' and 'future' observations are:

$$\hat{p}^{-}(t) = E\{p(t)|\{dn(\theta), 0 \leq \theta < t\}\}$$
$$\hat{p}^{+}(t) = E\{p(t)|\{dn(\theta), t \leq \theta < T\}\} \qquad (8)$$

We will first concentrate on the 'past observations' likelihood ratio, $L^{-}(t)$, that is a function of $\hat{p}^{-}(t)$ according to (7). A differential equation to obtain $\hat{p}^{-}(t)$ has been derived by Van Schuppen (1977:theorem 4.2), see also Segall et al (1975), Vaca and Snyder (1978) and Snyder (1975). It reads:

$$d\hat{p}^{-}(\theta) = \{[1-\hat{p}^{-}(\theta)]/\tau_0 - \hat{p}^{-}(\theta)/\tau_1\}d\theta + \qquad (9)$$
$$+ \frac{(r_1-r_0)\cdot\hat{p}^{-}(\theta)\cdot[1-\hat{p}^{-}(\theta)]}{r_0+(r_1-r_0)\cdot\hat{p}^{-}(\theta)}\cdot\{dn(\theta)-[r_0+(r_1-r_0)\cdot\hat{p}^{-}(\theta)]\cdot d\theta\}$$

Driving (9) from θ=0 to θ=t yields $\hat{p}^{-}(t)$. The initial condition, $\hat{p}^{-}(0)$, depends on experimental conditions. For instance, if we are rather sure that the monitor is started during wakefulness, $\hat{p}^{-}(0)\simeq 1$. Equations (9) and (7) are sufficient for the recursive computation of $L^{-}(t)$. However, the number of required multiplications can be reduced and the algorithm can be interpreted more clearly by the application of a logarithmic transformation. For notational simplicity we will use the ln rather than the ^{2}log from our earlier paper (Kemp et al 1985). The transformed variable equals:

$$\lambda(t) = \ln\{L(t)\} = \ln\{L^{-}(t)\}+\ln\{L^{+}(t)\} = \lambda^{-}(t)+\lambda^{+}(t) \qquad (10)$$

where, according to (7):

$$\lambda^{-}(t)= \ln\{(\tau_0/\tau_1)\cdot\hat{p}^{-}(t)/[1-\hat{p}^{-}(t)]\} \qquad (11)$$

The differential equation for the computation of $\lambda^{-}(t)$ can be obtained from (11) and (9). However, because (9) is driven by discontinuous counts, this may not be done by the ordinary chain rule for differentiation, nor by Itô's differentiation rule, but the Itô/Doléans-Dade/Meyer differentiation rule that accounts for these discontinuities (Van Schuppen 1977, Boel et al 1975, Vaca and Snyder 1978, Segall et al 1975, Segall 1976) must be applied as follows:

$$d\lambda^{-}(\theta) = [\delta\lambda^{-}(\theta)/\delta\hat{p}^{-}(\theta)]\cdot\omega\{d\hat{p}^{-}(\theta)\} + \Delta\{\lambda^{-}(\theta)\}\cdot dn(\theta) \qquad (12)$$

where $\delta./\delta.$ denotes a partial derivative, $\omega\{d\hat{p}^{-}(\theta)\}$ denotes the continuous part (i.e. $dn(\theta)=0$) of $d\hat{p}^{-}(\theta)$ in (9) and $\Delta\{\lambda^{-}(\theta)\}$ denotes the discontinuous jump in $\lambda^{-}(\theta)$ that is caused by a count, $dn(\theta)=1$. According to (11), this jump equals:

$$\Delta\{\lambda^{-}(\theta)\} = \ln\{(\tau_0/\tau_1)\cdot[\hat{p}^{-}(\theta)+\Delta\{\hat{p}^{-}(\theta)\}]/[1-\hat{p}^{-}(\theta)-\Delta\{\hat{p}^{-}(\theta)\}]\} - \ln\{(\tau_0/\tau_1)\cdot\hat{p}^{-}(\theta)/[1-\hat{p}^{-}(\theta)]\} \qquad (13)$$

where $\Delta\{\hat{p}^{-}(\theta)\}$ denotes the jump in $\hat{p}^{-}(\theta)$ that is caused by a count, $dn(\theta)=1$. According to (9), this jump equals:

$$\Delta\{\hat{p}\ (\theta)\} = (r_1-r_0)\cdot\hat{p}^{-}(\theta)\cdot[1-\hat{p}^{-}(\theta)] \ / \ \{r_0+(r_1-r_0)\cdot\hat{p}^{-}(\theta)\} \qquad (14)$$

Substitution of (14) in (13) shows that:

$$\Delta\{\lambda^{-}(\theta)\} = \ln(r_1/r_0) \qquad (15)$$

The partial derivative in (12) can be obtained from (11):

$$\delta\lambda^{-}(\theta)/\delta\hat{p}^{-}(\theta) = 1/\{\hat{p}^{-}(\theta)\cdot[1-\hat{p}^{-}(\theta)]\} \qquad (16)$$

Substitution of (15), (16) and the continuous part, $\omega\{d\hat{p}^{-}(\theta)\}$, of (9) in (12) yields the differential equation for $\lambda^{-}(t)$:

$$d\lambda^{-}(\theta)=\{[e^{-\lambda^{-}(\theta)}-1]/\tau_1-[e^{\lambda^{-}(\theta)}-1]/\tau_0\}d\theta + \ln(r_1/r_0)\cdot dn(\theta)-(r_1-r_0)d\theta \qquad (17)$$

Driving (17) from θ=0 to θ=t yields $\lambda^{-}(t)$. The initial condition, $\lambda^{-}(0)$ is a function (11) of $\hat{p}^{-}(0)$, which depends on the state of the subject when the monitor is started.

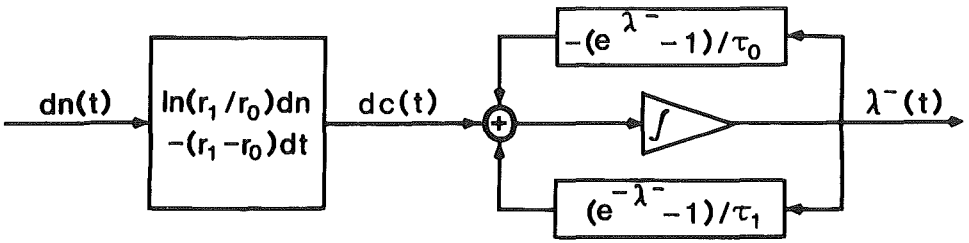


Figure 3: Block diagram of the exponential feedback integrator that generates the 'past observations' test statistic, $\lambda^{-}(t)$, if driven by the observed REM-bursts, dn(t).

Figure 3 shows a block diagram of the algorithm. It appears to be an integrator with exponential feedback that is driven by:

$$dc(\theta) = \ln(r_1/r_0)\cdot dn(\theta) - (r_1-r_0)\cdot d\theta \qquad (18)$$

The interpretation of this driving term is simple. Assume $r_1>r_0$. Between REM-bursts, i.e. $dn(\theta)=0$, the integrator will be driven to negative values while REM-bursts kick it to more positive values. According to (2) and (3), the conditional expectation:

$$E\{dn(\theta)|p(\theta)\} = \{r_0+(r_1-r_0)\cdot p(\theta)\}\cdot d\theta \qquad (19)$$

which implies that the conditional expectation of the driving term equals:

$$E\{dc(\theta)|p(\theta)\} = \{\{r_0+(r_1-r_0)\cdot p(\theta)\}\cdot \ln(r_1/r_0) - (r_1-r_0)\}\cdot d\theta$$
$$\overset{6}{\underbrace{\quad}}\ \kappa_f\cdot\{(r_1/r_0)^{p(\theta)}\cdot \ln(r_1/r_0) - r_1/r_0 + 1\}\cdot d\theta \qquad (20)$$

error by Océ printer driver, correct in article Biol. Cyb. Bd.

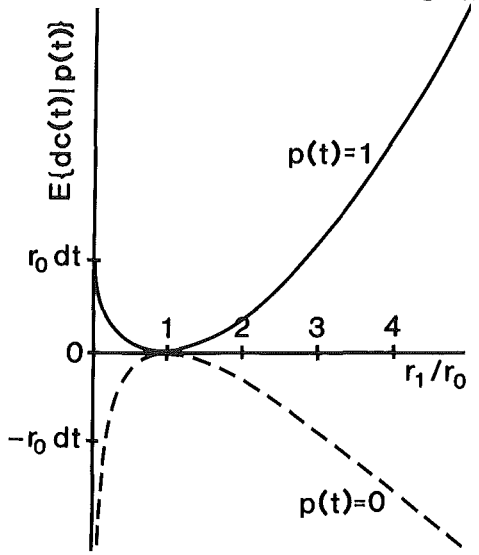


Figure 4: Relationship between the conditional expectation, $E\{dc(t)|p(t)\}$, of the driving term, $dc(t)$, and the REM-rate ratio, r_1/r_0. Note that, for any ratio, the sign of the expectation directly corresponds to $p(t)$.

Figure 4 shows the behaviour of this term for different values of (r_1/r_0) and of $p(\theta)$. Apparently, the integrator is, on average, driven to positive values during REM-states (i.e. $p(\theta)=1$) and to negative values during NREM-states (i.e. $p(\theta)=0$). The saturation effect that is caused by the exponential feedback limits the time needed to react to a transition of $p(\theta)$. Quite satisfactory, this effect increases with increasing transition rates.

Because (1) and (2) may also describe the system in the backward time direction, provided we invert the sign of the Poisson counts, $dn(\theta)$, and of

the time increments, $d\theta$, the transformed 'future' test statistic, $\lambda^+(t)$ follows from the corresponding differential equation:

$$d\lambda^+(\theta) = \{[e^{-\lambda^+(\theta)}-1]/\tau_1 - [e^{\lambda^+(\theta)}-1]/\tau_0\}\cdot\{-d\theta\} + \\ + \ln(r_1/r_0)\cdot\{-dn(\theta)\}-(r_1-r_0)\cdot\{-d\theta\} \qquad (21)$$

except that this is a backward differential equation that has to be driven from θ=T ('future') to θ=t ('present') and consequently by non-positive counts, $dn(\theta)$, and time increments, $d\theta$, in order to obtain $\lambda^+(t)$. The initial condition, $\lambda^+(T)$, is a function of $\hat{p}^+(T)$ and depends on the state of the subject at the end of the registration.

Because the likelihood ratio is a function of $\lambda^-(t)$ and $\lambda^+(t)$:

$$L(t) = e^{\lambda^-(t)+\lambda^+(t)} \qquad (22)$$

the problem is now solved. However, although (17) and (21) enable one to understand and provide a plausible interpretation of the algorithm, it can be transformed into a computationally more efficient form. By substitution of (22), both the forward equation (17, i=-) and the backward equation (21, i=+) can be written as a differential equation for $L_i(\theta)$:

$$dL^i(\theta) = \{[1-L^i(\theta)]/\tau_1 - [L^{i^2}(\theta)-L^i(\theta)]/\tau_0 - L^i(\theta)\cdot[r_1-r_0]\}\cdot\{-i\cdot d\theta\} + \\ + L^i(\theta)\cdot\ln(r_1/r_0)\cdot\{-i\cdot dn(\theta)\} \qquad i\varepsilon\{-,+\} \qquad (23)$$

Between counts, i.e. $dn(\theta)$=0, this is a separable first order differential equation with the analytical solution which allows direct computation of $L^i(\theta)$ from $L^i(\phi)$:

$$L^i(\theta) = \frac{\rho_1 \cdot [\rho_2 - L^i(\phi)] - \rho_2 \cdot [\rho_1 - L^i(\phi)] \cdot e^{i \cdot (\phi - \theta)/a}}{[\rho_2 - L^i(\phi)] - [\rho_1 - L^i(\phi)] \cdot e^{i \cdot (\phi - \theta)/a}} \qquad (24)$$

where $i \in \{-,+\}$ and $i(\phi-\theta) \geq 0$. The constants are $\rho_1 = \frac{1}{2}\{-b+\sqrt{c}\}$, $\rho_2 = \frac{1}{2}\{-b-\sqrt{c}\}$, and $a = -\tau_0/\sqrt{c}$, where $b = \tau_0/\tau_1 - 1 + \tau_0 \cdot (r_1 - r_0)$ and $c = b^2 + 4\tau_0/\tau_1$. According to (17) and (21) a count adds $\ln(r_1/r_0)$ to $\lambda^i(\theta)$. Therefore, according to (22), a count causes $L^i(\theta)$ to be multiplied by r_1/r_0:

$$L^i(\theta + d\theta) = L^i(\theta) \cdot (r_1/r_0)^{-i \cdot dn(\theta)} \qquad (25)$$

We computed $L^-(t)$ forward (starting at $\theta=0$) and $L^+(t)$ backward (starting at $\theta=T$) with a time resolution of 1s, according to (24) between counts and according to (25) at counts. Multiplication according to (6) yields the likelihood ratio, L(t), that was compared to the threshold, K, in order to obtain the optimal decision, $\overset{*}{p}(t)$, on p(t).

IV.4. Performance

A rather simple performance analysis was carried out by simulations using our REM-burst generating model. Although the obtained values are not necessarily valid for real sleep, some interesting qualitative conclusions can be drawn.

Both the brain state process, p(t), and the corresponding observations, dn(t), have been simulated according to (1), (2) and (3) and the corresponding exponential distributions (method in Snyder 1975:p.62). The resulting observations of 100 'nights' of 8 hours, all starting in the awake state, were then fed into the monitor that provides the decision, $\overset{*}{p}(t)$, on p(t). We used initial conditions:

$L^-(0) \simeq \infty$ (i.e. the largest number our computer could handle)

$$L^+(T) = 1 \qquad (26)$$

that can be derived by (7) from the initial estimates:

$\hat{p}^-(0) = 1$, because the monitor starts in the awake state.

$$\hat{p}^+(T) = E\{p(\infty)\} = \tau_1/(\tau_0+\tau_1) \qquad (27)$$

, assuming that the stationary situation has been reached at the end of the night.

monitor parameters				model parameters	performance		
τ_1	τ_0	r_1	r_0		P(1|1)	P(0|0)	P(i|i)
10	30	.5	.04		.749	.953	.899
20	60	1	.02		.949	.985	.975
40	120	2	.01		.990	.997	.995
20	60	1	.02	τ_1=4	.769	.989	.974
20	60	1	.02	τ_0=12	.953	.897	.931
20	60	1	.02	r_1=.2	.170	.998	.773
20	60	1	.02	r_0=.1	.953	.932	.938
20	60	1	.02	τ_1=100	.983	.984	.983
20	60	1	.02	τ_0=300	.953	.997	.993
20	60	1	.02	r_1=5	.999	.971	.979
20	60	1	.02	r_0=.004	.937	.992	.978

Table 1. Performance of the monitor for different sojourn times, τ_1 and τ_0 (minutes), and rates, r_1 and r_0 ($(min)^{-1}$). Model parameters are listed only if they differ from the corresponding monitor parameters. P(1|1), P(0|0) and P(i|i) are the probabilities of correct classification of the REM-state, the NREM-state and the state, respectively. Note that sojourn times, differing by a factor 5 from the standard values of figure 2, hardly decrease performance.

An example is given in figure 2. By comparing $\overset{*}{p}(t)$ to p(t) during the 800 hour simulation, the probability of a correct decision, $P\{\overset{*}{p}(t)=p(t)\}$, was estimated. For an 'average patient', having REM rates r_1=1/minute and r_0=.02/minute and average sojourn times τ_0=60 minutes and τ_1=20 minutes, this probability is 97%.

The results for other rates and sojourn times, including some in which the model parameters differ from the monitor parameters, have been summarized in table 1.

IV.5. Discussion

As in our earlier paper (Kemp et al 1985), the optimal temporal smoother appears to be an integrator with exponential feedback that depends only on the brain state statistics, τ_0 and τ_1. Because we are now dealing with discontinuous point instead of continuous observations, the driving term is different. It is a rather unexpected, but effective (fig.4) function of the REM-burst rates r_0 and r_1.

For practical sleep stage scoring we conclude that, using mathematical models, original and attractive solutions may be created for the processing of both continuous and discontinuous observations. In particular, the exponential feedback integrator might prove to be a simple and effective alternative to the usually applied segmentation of sleep recordings into short, stationary, intervals. Also, the rather erratic statistical behaviour of the REM-bursts and of the brain state does not necessarily abolish the possibility of obtaining 97% correct decisions. And finally, monitoring performance is hardly degraded by the small mistunings of these statistics which will inevitably occur in practice (table 1).

Acknowledgements. We gratefully acknowledge the constructive criticism and help of Dr. J.H. Van Schuppen of the Centre for Mathematics and Computer Science in Amsterdam, Professor Ir. E.W. Gröneveld of the Twente University of Technology and Professor Dr. F.H. Lopes Da Silva of the University of Amsterdam.

V. A MODEL-BASED MONITOR OF HUMAN SLEEP STAGES

B. Kemp, E.W. Gröneveld, A.J.M.W. Janssen, J.M. Franzen

Submitted for publication
Published in: Biological Cybernetics 1987, 57: 365-378

Summary

Stochastic models are proposed for sleep and for the sleep related electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG). The evolution of sleep through its various stages is described as a Markov chain. The EEG is modelled using Wiener processes. The EOG and EMG are modelled as combinations of Poisson point processes and Gaussian processes, respectively. The EEG models contain a feedback structure that is based on physiological data. The maximum likelihood sleep stage monitor, that uses the sleep-related observations, has been derived and implemented. The agreement between automatic and human stage classifications of six sleep recordings was 70.6%, which was 4.5% worse than the average agreement between six human classifiers. Monitoring of simulated sleep suggests that the difficulty in separating wakefulness from stage 1 is due to poor modelling. If one ignores this difference, which, from a diagnostic point of view is fairly unimportant, the above mentioned agreement reaches 81.8%, which is 0.5% better than the corresponding average human vs human agreement.

V.1. Introduction

In the course of a night human sleep travels through various stages. These stages correspond to (a.o.) typical characteristics of the electrical activity of the brain, eye movements and muscle tension (Rechtschaffen and Kales 1968). Sleep classification (fig.1) is therefore generally performed by screening the recorded electroencephalogram (EEG), electrooculogram (EOG) and

electromyogram (EMG) signals for these characteristics. If carried out in this way, visual reading and subjective analysis of a whole-night record take about 6 hours. Usually, different human classifiers obtain rather inconsistent results (Johnson 1977, Monroe 1969).

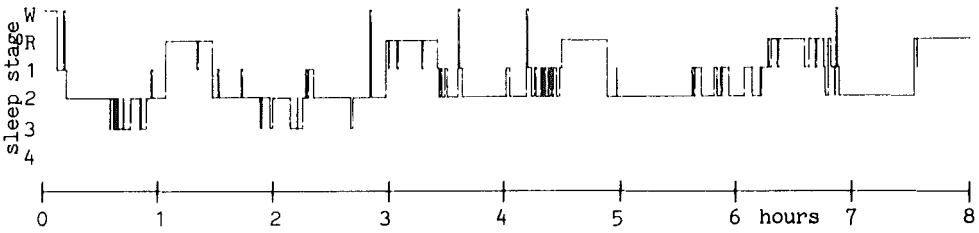


Figure 1: Classification of one night's sleep into stages Wakefulness, Rem, 1, 2, 3 and 4.

Automatic sleep stage monitoring (Smith et al 1978 (overview), Lim and Winters 1980, Gath and Bar-On 1980, Inoue et al 1982, Hoffmann et al 1984, Chouvet et al 1980, Hermann and Kubicki 1984, Johnson 1977, Broughton et al 1979, Jansen et al 1980, Lacroix and Stanus 1985[b], Kumar 1977, Hasan 1983, Hasan 1985 (review), Campbell and Wilkinson 1981, Martens et al 1982, Smith 1986) is objective and less laborious. However, in the studies available, the agreement between automatic and human classifications is smaller than the agreement between human observers (Johnson 1977, Lacroix and Stanus 1985[a,b]). This suggests that automatic procedures may still be improved. Automatic systems have grown increasingly complex, so much so that they have hardly been reported in sufficient detail to allow reproduction and interpretation. In all systems, some Boolean matrix or decision tree is applied to preliminary decisions on isolated EEG, EOG and EMG characteristics, like the presence or absence of an alpha rhythm in the EEG. No systematic approach to the optimal design of such a tree exists. This use of intermediate decisions in the process of obtaining one final decision considerably reduces the infor-

mation from the original signals and is therefore a suboptimal procedure (Woodward 1953:ch.3.7). We have shown (Kemp et al 1985, Kemp 1986) that original and rather simple signal processors can be derived, using likelihood ratios, from models of sleep and some of the sleep-related signals. Because the models are partly based on physiological data, the results may be more closely related to physiological reality than the commonly applied methods like Fourier analysis or ARMA filtering, that are imported from technical environments. The likelihood ratio method will, by definition, combine the various EEG, EOG and EMG characteristics in such a way that no information is lost.

These arguments have led us to develop a comprehensive model of sleep and sleep-related EEG, EOG, and EMG observations. The maximum-likelihood sleep stage monitor that corresponds to this model will be derived. Its performance on both simulated (by the model) and on real sleep will be demonstrated.

V.2. Problem formulation: the model

In this section, we propose models for sleep itself and for the sleep-related EEG, EOG, and EMG signals. The problem of how to monitor sleep using these observed signals will then be formulated in terms of the model.

We will restrict our observation models to the sigma-, alpha-, and delta-rhythms in the EEG, the very rapid and the very slow eye-movements (REMs and SEMs, respectively) in the EOG and finally the submental muscle tone that can be obtained from the EMG. This choice is based on our experience that these signal characteristics show a similar sleep dependency in most subjects and can be detected fairly reliably. Most investigators, involved in automatic sleep stage monitoring apply these characteristics. In some cases, theta- or beta-rhythms, K-complexes, vertex sharp waves, sawtooth

waves, etc. are also considered. The models and the mathematical derivations in this paper can be generalized to include such features.

V.2.1. The sleep process

We conform to the common proposition in sleep research (e.g. Rechtschaffen and Kales 1968) that in the time, t, course of a night, $0 \leq t < T$, the sleep state, h(t), that is to be monitored, can be classified into a set, $S=\{W,R,1,2,3,4,\}$ of H=6 stages; Wakefulness, Rem-sleep and the sleep stages 1, 2, 3, and 4. The transitions between these stages occur in unpredictable directions and at random points in time (fig.1). Recently (Kemp and Kamphuisen 1986), we proposed a model that accounts for most deterministic and statistical aspects of sleep stage patterns. It is an inhomogeneous, continuous-time, Markov chain (e.g. Larson and Shubert 1979:ch.3.3). Such a model is characterized by time-varying transition rates,

$$a_{j|i}(t) = \lim_{\delta \to 0} [P\{h(t+\delta)=j | h(t)=i\}/\delta] \qquad i \neq j \qquad i,j \in S \qquad (1)$$

where $P\{.|.\}$ denotes a conditional probability. The time dependencies of various rates reflect the well known periodicities and trends in sleep, that are different for each individual. The Markov chain is asymmetric in the state space, i.e. the rates between different stages differ because some transitions occur more frequently and some sleep stages last longer than others. However, reliable a priori information about these individual dynamics and asymmetries is generally not available, particularly when sleep pathology is involved. Therefore, for our present purpose, we simplify the model to a homogeneous and symmetric, continuous-time Markov chain, i.e. with constant and equal transition rates:

$$a_{j|i}(t) = a \qquad i \neq j \qquad i,j \in S \qquad (2)$$

Another argument for this simplification is our intention to use the monitor for quantification of these dynamics and asymmetries; thus, this quantification should not be biased by a priori assumptions.

It is possible to derive continuous-time sleep state monitors, that are optimal for such models (Kemp and Jaspers 1984, Kemp et al 1985, Kemp 1986). However, the derivations were based on nonlinear filters that are not yet available for our case in which both continuous signals and point processes are observed. Also, the monitor must eventually be implemented on a digital computer and compared to the human classification that is based on segmentation in 30s epochs. Approximation in discrete time with interval, D (e.g. D=30s), is therefore necessary. This approximation may cause severe suboptimalities, e.g. instability (Clements and Anderson 1973). Having to accept this limited time resolution, we may as well reformulate the problem in discrete time, i.e. only the sleep state samples, h(nD), where n denotes the sample number, have to be monitored. We assume that the EEG, EOG and EMG generating mechanisms are modulated by these samples only. In this way, the sleep-dependencies are updated at the sample times, nD, whereas during the intervals, $[nD,nD+D)$, they remain constant. Besides the limited time resolution, D, no other suboptimalities are to be expected.

The homogeneous Markov property of h(t) implies (e.g. Larson and Shubert 1979:ex.3.3.3) that the samples, h(nD), form a discrete-time Markov chain with transition probabilities:

$$A = P\{h(nD+D)=j \mid h(nD)=i\} = P\{h(nD)=j \mid h(nD+D)=i\} = \{1-\exp(-H\cdot a\cdot D)\}/H \qquad (3)$$

$$i\neq j \qquad i,j\in S$$

For the remainder of this article we will assume n=0,1,2,..,N-1 unless specified otherwise.

V.2.2. EEG observations

We have proposed continuous-time generator models (fig.2) for both the alpha rhythm (Kemp and Blom 1981) and the sigma rhythm (Kemp et al 1985). These models are based on the assumptions that these rhythms originate from white-noise-driven feedback structures in which the frequency-selective feedback may be inhibited by 'external' influences like sleep. These assumptions are supported by physiological and histological observations on sigma- (Shouse and Sterman 1979, Sterman and Bowersox 1981), alpha- (e.g. Lopes da Silva et al 1976, Andersen et al 1966) and delta- (Corner 1984) rhythms.

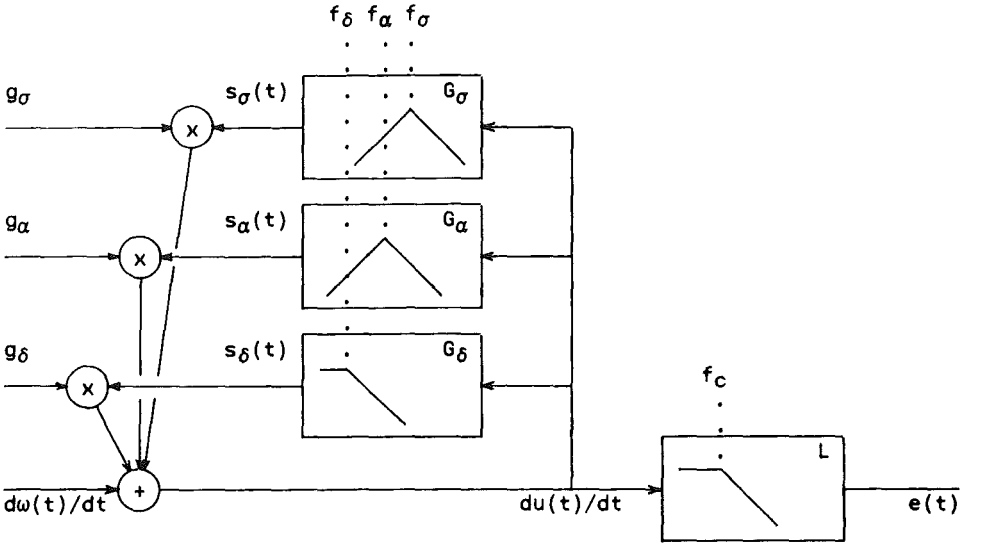


Figure 2: White noise ($d\omega(t)/dt$) driven structure with σ-, α- and δ- feedback, responsible for the corresponding components in the EEG, e(t). The transfer functions of the analogue linear filters are bandpasses, $G_\sigma(f)=1/[1+j\cdot(f_\sigma/B)\cdot(f/f_\sigma-f_\sigma/f)]$, $G_\alpha(f)=1/[1+j\cdot(f_\alpha/B)\cdot(f/f_\alpha-f_\alpha/f)]$, and lowpasses, $G_\delta(f)=1/[1+j\cdot(f/f_\delta)]$, $L(f)=1/[1+j\cdot(f/f_c)]$. The centre frequencies, f_σ and f_α, are tuned (using power spectral analysis) to the centre frequencies of the subjects σ- and α-rhythm, respectively. The bandwidths are B=3.5Hz and the cut-off frequencies are f_δ=1.9Hz and f_c=1.8Hz. L(f) simulates the volume conduction from cortical to scalp EEG. The feedback gains, $g_.(i)$ depend on the sleep stage, i (see table 1).

According to this model (fig.2) we can replace the scalp EEG observations, e(t), by the cortical EEG observations, du(t), since these can be reconstructed from e(t) by the inverse, $L^{-1}(f)$, of the volume conduction filter and vice versa. From figure 2 we obtain the observations equation:

$$du(t) = g_{\sigma}(h(nD))\cdot s_{\sigma}(t)\cdot dt + g_{\alpha}(h(nD))\cdot s_{\alpha}(t)\cdot dt + g_{\delta}(h(nD))\cdot s_{\delta}(t)\cdot dt + d\omega(t) \qquad nD \leq t < nD+D \qquad (4)$$

where $\omega(t)$ is the standard Wiener process (i.e. $d\omega(t)/dt$ is standard Gaussian white noise), the feedback signals, $s_{\cdot}(t)$ can be obtained on-line (by filtering by $G_{\cdot}(f)$, assuming the initial state of the feedback filter is known) from du(t). Finite conduction velocities always cause small time delays in physiological systems. Thereby the increment, $d\omega(t)$, is independent of the feedback signals, $s_{\cdot}(t)$. Accordingly, the increments, $d\omega(t)$ and du(t), should be interpreted as forward increments with respect to $s_{\cdot}(t)$. In this way, the independence property is maintained in the model.

The feedback gains, $g_{\cdot}(.)$, are modulated by the samples, h(nD), of the sleep process, h(t). Sigma rhythms are present mainly at fronto-central locations and alpha rhythms mainly at the back of the skull, e.g. on Fpz-Cz and Pz-Oz electrode locations (Jasper 1958), respectively. We have, therefore, modelled two corresponding observations, $du_F(t)$ and $du_B(t)$, in which $g_{\alpha}(h(nD))=0$ and $g_{\sigma}(h(nD))=0$, respectively. Because the low-frequency component, $g_{\delta}(h(nD)\cdot s_{\delta}(t)\cdot dt$, of $du_F(t)$ will be considerably disturbed by eye blinks, the monitor should not use it. Therefore, in $du_F(t)$, we set $g_{\delta}(h(nD))=0$. The resulting two EEG models are (with $nD \leq t < nD+D$):

$$du_F(t) = g_{\sigma}(h(nD))\cdot s_{\sigma}(t)\cdot dt + d\omega_F(t) \qquad (5)$$

$$du_B(t) = g_{\alpha}(h(nD))\cdot s_{\alpha}(t)\cdot dt + g_{\delta}(h(nD))\cdot s_{\delta}(t)\cdot dt + d\omega_B(t) \qquad (6)$$

V.2.3. Eye movement observations

A variety of EOG preprocessors for the detection of REM-bursts and SEMs have been described (e.g. McPartland and Kupfer 1978, Ktonas and Smith 1978, Toth 1971, Goldberg and Beiber 1979, Gopal and Haddad 1981, Kayed et al 1979, Lacroix and Stanus 1985[a,b], Smith 1978). We recorded the horizontal eye movements using electrodes placed at the outer canthus of each eye. REM-bursts and SEMs were detected from this EOG using bandpass and lowpass filtering, respectively. Recently (Kemp 1986), we proposed a model that describes the statistics of the detected REM-bursts. It is a Poisson counting process, $u_R(t)$, with sleep-dependent rate. Each count, $du_R(t)=1$, corresponds to a REM-burst and between REM-bursts $du_R(t)=0$. In short:

$$u_R(t) \text{ is a Poisson-process with rate, } r_R(h(nD)) \qquad nD \leq t < nD+D \qquad (7)$$

Because short intervals between successive SEMs occur more frequently than longer ones, we propose another sleep-dependent Poisson counting process, $u_S(t)$, for the SEMs:

$$u_S(t) \text{ is a Poisson process with rate, } r_S(h(nD)) \qquad nD \leq t < nD+D \qquad (8)$$

Each count, $du_S(t)=1$, corresponds to a detected SEM and between SEMs $du_S(t)=0$.

V.2.4. EMG observations

Several methods of measuring the muscle tone from the recorded EMG have been reported (e.g. Inoue et al 1982, Othmer and Othmer 1980, Hasan 1983, Lacroix and Stanus 1985[a]). We have adopted the commonly applied rectification and averaging of the submental EMG during segments, d=1s. The resulting discrete-time observations, $u_M(md)$ with $md=0,d,2d,\ldots,ND-d$, are assumed to be

mutually conditionally (on h(nD)) independent because the bandwidth of the EMG is much larger than 1Hz. Histograms (fig.3) support our assumption that the probability law of $u_M(md)$ may be approximated by a linear combination of two Gaussian distributions, at least in our (low-tone) region of interest. We also know by experience (see also Rechtschaffen and Kales 1968) that muscle tone during REM-sleep is low, except for short (about ½ to 3s) twitches that seem to occur at random points in time. We therefore assume that the samples, $u_M(md)$, are selected mutually independently and with sleep-dependent probability, $q_M(h(nD))$, from the high-tone Gaussian $N(\mu_1,\sigma_1^2)$ distribution. The probability for the low-tone $N(\mu_0,\sigma_0^2)$ distribution is, of course, $(1-q_M(h(nD)))$. For example, during wakefulness, muscle tone is always high, i.e. $q_M(W)=1$.

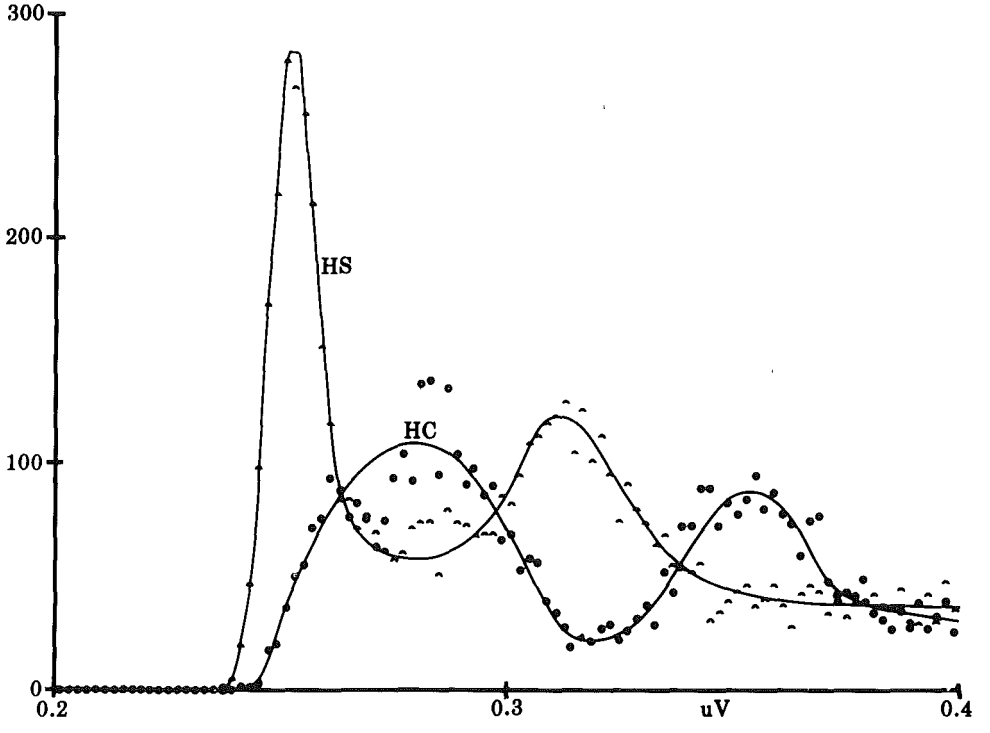


Figure 3: Whole-record (2½h) histogram of the muscle tone $u_M(md)$, in the two subjects (HS and HC) of the learning set (see section V.4). Note that both subjects show both peaks, μ_0 and μ_1, in the low-tone region.

Summarizing:

$$u_M(md) \text{ are } \{1-q_M(h(nD))\}\cdot N(\mu_0,\sigma_0^2) + q_M(h(nD))\cdot N(\mu_1,\sigma_1^2) \text{ distributed} \quad (9)$$

in which $nD \leq md < nD+D$. The parameters, μ_0 and μ_1, must be adapted to the subject by detecting the two peaks in the histogram of figure 3. In all subjects we chose $\sigma_0=\sigma_1=\mu_1-\mu_0$.

V.2.5. Problem formulation

The models (5)-(9) describe the mechanisms by which the observations are influenced by the sleep stages. Because the observed signals arise from different brain structures, we assume them to be mutually independent except for the common modulation of some parameters by h(nD). The sleep-dependency of these parameters can be summarized in the form of a sleep characteristics matrix (table 1).

		observations characteristics					
	i	$g_\sigma(i)$	$g_\alpha(i)$	$g_\delta(i)$	$r_R(i)$	$r_S(i)$	$q_M(i)$
sleep stage	W	.2	.7	.40	1.0	1.2	1.
	R	.2	.7	.40	1.0	1.2	.2
	1	.2	.7	.40	.05	6.0	1.
	2	.5	.5	.60	.05	1.2	1.
	3	.5	.5	.85	.05	1.2	1.
	4	.5	.5	.98	.05	1.2	1.

Table 1: Typical sleep characteristics matrix, describing the relationship between the sleep stage, i, and the sleep dependent sigma-, alpha-, and delta-feedback gains $g_\sigma(i)$, $g_\alpha(i)$ and $g_\delta(i)$ in the EEG, the rapid- and slow-eye-movement rates (min^{-1}), $r_R(i)$ and $r_S(i)$, in the EOG and the high-muscle-tone probability, $q_M(i)$ of the EMG.

Based on this description (fig.4) of the sleep stage, $h(nD) \in S$, and its relationship to the observations, the problem can be formulated as follows. Which monitor provides at any discrete time, nD, during the recording interval, $0 \leq t < ND$, the optimal decision, $\hat{h}(nD)$, based on the full observations realization, $U(0,ND)$, consisting of the X=5 scalar observations (5)-(9):

$$U(0,ND) = \begin{bmatrix} U_F(0,ND) \\ U_B(0,ND) \\ U_R(0,ND) \\ U_S(0,ND) \\ U_M(0,ND) \end{bmatrix} = \begin{bmatrix} du_F(t) , & 0 \leq t < ND \\ du_B(t) , & 0 \leq t < ND \\ du_R(t) , & 0 \leq t < ND \\ du_S(t) , & 0 \leq t < ND \\ u_M(md) , & md = 0,d,2d,..ND-d \end{bmatrix} \qquad (10)$$

Note that a decision at time, nD, is valid during the interval, $[nD,nD+D)$. As an optimization rule, we have adopted the maximum a posteriori probability criterion (e.g. Van Trees 1979).

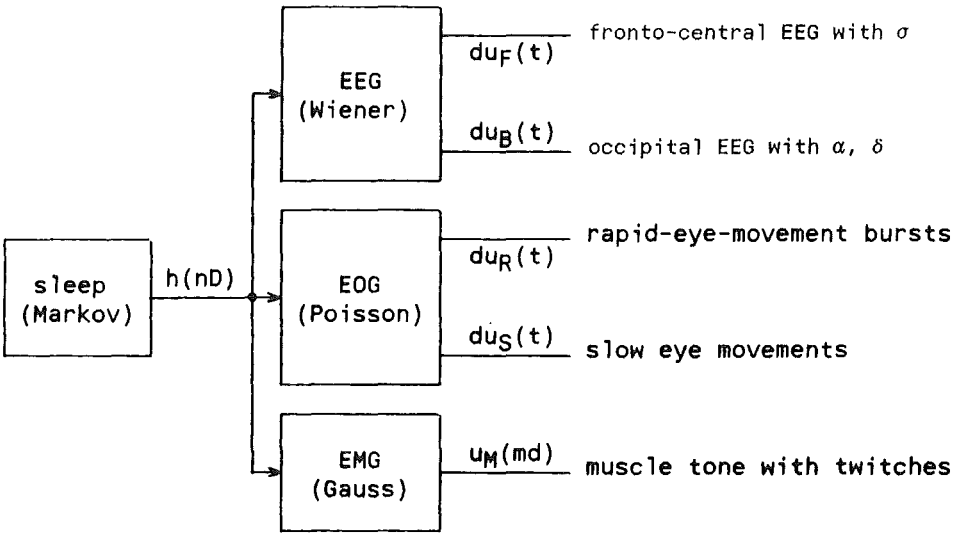


Figure 4: Abstract model of sleep and sleep-related observations.

V.3. Derivation of the optimal monitor

The decision, $\overset{*}{h}(nD)=j$, that is most probably correct, is the one that maximizes (over $i\epsilon S$) the conditional a posteriori stage probability:

$$p_i(nD) = P\{h(nD)=i \mid U(0,ND)\} \tag{11}$$

The probability that this decision is correct equals $p_j(nD)$. Because the Markov chain is homogeneous (in time and in space) with H states, the a priori probabilities are:

$$P\{h(nD)=i\} = 1/H \tag{12}$$

Throughout this section we will assume $i\epsilon S$. Using (11), (12) and Bayes' rule we can write $p_i(nD)$ as a function of conditional probability densities, $f\{.|.\}$, of the observations as follows:

$$p_i(nD) = p_i(nD)/\sum_{j\epsilon S} p_j(nD) = f_i(nD)/\sum_{j\epsilon S} f_j(nD) \tag{13}$$

$$\text{in which } f_i(nD) = f\{U(0,ND) \mid h(nD)=i\} \tag{13a}$$

which implies that maximizing $p_i(nD)$ is equivalent to maximizing (also over $i\epsilon S$) the full observations probability density, or likelihood, $f_i(nD)$. These H densities can be factorized into a 'past', 'present' and 'future' observations part, respectively:

$$f_i(nD) = f_i^-(nD)\cdot f_i^o(nD)\cdot f_i^+(nD) \quad \text{, in which} \tag{14}$$

$$f_i^-(nD) = f\{U(0,nD) \mid h(nD)=i\} \tag{14a}$$

$$f_i^o(nD) = f\{U(nD,nD+D) \mid U(0,nD),h(nD)=i\} \tag{14b}$$

$$f_i^+(nD) = f\{U(nD+D,ND) \mid U(0,nD+D),h(nD)=i\} \tag{14c}$$

The past observations parts, $f_i^-(nD)$, can be computed by a recursive equation that is driven by $f_i^o(nD)$ as follows (n=1,2,...,N-1):

$$
\begin{aligned}
f_i^-(nD) &= f\{U(0,nD)\,|\,h(nD)=i\} = \\
&= \sum_{j\epsilon S} \left[P\{h(nD-D)=j\,|\,h(nD)=i\}\cdot f\{U(0,nD)\,|\,h(nD)=i,h(nD-D)=j\}\right] = \\
&= \sum_{j\epsilon S} \left[P\{h(nD-D)=j\,|\,h(nD)=i\}\cdot f\{U(0,nD)\,|\,h(nD-D)=j\}\right] = \\
&= \sum_{j\epsilon S} \left[P\{h(nD-D)=j\,|\,h(nD)=i\}\cdot f_j^-(nD-D)\cdot f_j^o(nD-D)\right] = \\
&= A\cdot\Big\{\sum_{j\neq i} \left[f_j^-(nD-D)\cdot f_j^o(nD-D)\right]\Big\} + \{1-(H-1)\cdot A\}\cdot f_i^-(nD-D)\cdot f_i^o(nD-D) = \\
&= A\cdot\Big\{\sum_{j\epsilon S} \left[f_j^-(nD-D)\cdot f_j^o(nD-D)\right]\Big\} + \{1-HA\}\cdot f_i^-(nD-D)\cdot f_i^o(nD-D) \qquad (15)
\end{aligned}
$$

with inital values $f_i^-(0)=1$. A similar recursive equation for the 'future' observations parts, $f_i^+(nD)$, runs in the backward time direction as follows (n=0,1,2,...,N-2):

$$
\begin{aligned}
f_i^+(nD) &= f\{U(nD+D,ND)\,|\,U(0,nD+D),h(nD)=i\} = \\
&= \sum_{j\epsilon S} \Big[P\{h(nD+D)=j\,|\,U(0,nD+D),h(nD)=i\}\cdot \\
&\qquad \cdot f\{U(nD+D,ND)\,|\,U(0,nD+D),h(nD)=i,h(nD+D)=j\}\Big] = \\
&= \sum_{j\epsilon S} \left[P\{h(nD+D)=j\,|\,h(nD)=i\}\cdot f\{U(nD+D,ND)\,|\,U(0,nD+D),h(nD+D)=j\}\right] \\
&= \sum_{j\epsilon S} \left[P\{h(nD+D)=j\,|\,h(nD)=i\}\cdot f_j^+(nD+D)\cdot f_j^o(nD+D)\right] = \\
&= A\cdot\Big\{\sum_{j\neq i} \left[f_j^+(nD+D)\cdot f_j^o(nD+D)\right]\Big\} + \{1-(H-1)A\}\cdot f_i^+(nD+D)\cdot f_i^o(nD+D) = \\
&= A\cdot\Big\{\sum_{j\epsilon S} \left[f_j^+(nD+D)\cdot f_j^o(nD+D)\right]\Big\} + \{1-HA\}\cdot f_i^+(nD+D)\cdot f_i^o(nD+D) \qquad (16)
\end{aligned}
$$

with initial values $f_i^+(ND-D)=1$.

Apparently, the full observations probability density $f_i(nD)$ can be computed using (14), (15) and (16) from the sequence of present observations probability densities, $f_i^o(nD)$. The latter are functions, (14^b), of the observations vector, U(nD,nD+D), that consists of the five scalar observations, $U_x(nD,nD+D)$ with $x\epsilon\chi=\{F,B,R,S,M\}$, as described in (10). Because these are mutually conditionally (on h(nD) and the 'past' observations) independent, each vector observations probability density, $f_i^o(nD)$, can be factorized into X=5 scalar observations probability densities as follows:

$$f_i^o(nD) = \prod_{x\epsilon\chi} f_{i,x}^o(nD) \quad \text{, in which} \tag{17}$$

$$f_{i,x}^o(nD) = f\{U_x(nD,nD+D) \mid U_x(0,nD), h(nD)=i\} \tag{17^a}$$

The probability densities, $f_{i,x}^o(nD)$, equal likelihoods that have been derived in the literature mentioned below, except for a proportionality factor that does not depend on i. Because this factor is the same in both numerator and denominator of (13), it may, and will, be neglected.

According to the EEG observations models (5) and (6), $f_{i,F}^o(nD)$ and $f_{i,B}^o(nD)$ then are likelihoods for random signals in white Gaussian noise. They read (e.g. Larson and Shubert 1979:ex.4.5.1, Kailath 1969):

$$f_{i,F}^o(nD) = \exp\{\int z_{i,F}(t)\cdot du_F(t) - \tfrac{1}{2}\int z_{i,F}^2(t)\cdot dt\} \tag{18}$$

$$f_{i,B}^o(nD) = \exp\{\int z_{i,B}(t)\cdot du_B(t) - \tfrac{1}{2}\int z_{i,B}^2(t)\cdot dt\} \tag{19}$$

in which the random signals are:

$$z_{i,F}(t) = g_\sigma(i)\cdot s_\sigma(t) \tag{18^a}$$

$$z_{i,B}(t) = g_\alpha(i)\cdot s_\alpha(t) + g_\delta(i)\cdot s_\delta(t) \tag{19^a}$$

The feedback signals, $s_{\cdot}(t)$, can be reconstructed by causal filtering (section 2.2) of the observations, $du_{\cdot}(t)$. Thus, although the feedback signals are not known a priori, they eventually appear as a priori conditions (see also eq. (17^a)) in the sufficient EEG-statistics, i.e. in the sequence of 'present' EEG likelihoods. Apparently, they can be assumed to be a priori known signals in the formulation of the problem as we did in our earlier article (Kemp et al 1985), based on intuitive reasoning.

All integrals in (18)-(21) are over the the interval $[nD,nD+D)$.

According to the EOG observations models (6) and (7), $f^o_{i,R}(nD)$ and $f^o_{i,S}(nD)$ are likelihoods for the rates, $r_R(i)$ and $r_S(i)$, of the Poisson counting processes, $u_R(t)$ and $u_S(t)$, respectively. They read (e.g. Snyder 1975:ex.2.5.2, Davis and Andreadakis 1977):

$$f^o_{i,R}(nD) = \exp\{\ln[r_R(i)]\cdot\int du_R(t) - r_R(i)\cdot\int dt\} \tag{20}$$

$$f^o_{i,S}(nD) = \exp\{\ln[r_S(i)]\cdot\int du_S(t) - r_S(i)\cdot\int dt\} \tag{21}$$

in which $\int du_R(t)$ and $\int du_S(t)$ are, according to (7) and (8), the number of observed REM-bursts and SEMs in the interval $[nD,nD+D)$, respectively. Of course, $\int dt=D$.

The likelihood, $f^o_{i,M}(nD)$, follows directly from the conditional Gaussian distributions of the EMG observations, $u_M(md)$. It reads:

$$f^o_{i,M}(nD) = \prod_{md=nD}^{nD+D-d} \{[1-q_M(i)]\cdot\exp\{-[u_M(md)-\mu_0]^2/2\sigma_0^2\}/\sigma_0 + \\ + q_M(i)\cdot\exp\{-[u_M(md)-\mu_1]^2/2\sigma_1^2\}/\sigma_1\} \tag{22}$$

All likelihoods, $f^o_{i,x}(nD)$, that are required in (17) can be computed directly from the observations using (18)-(22).

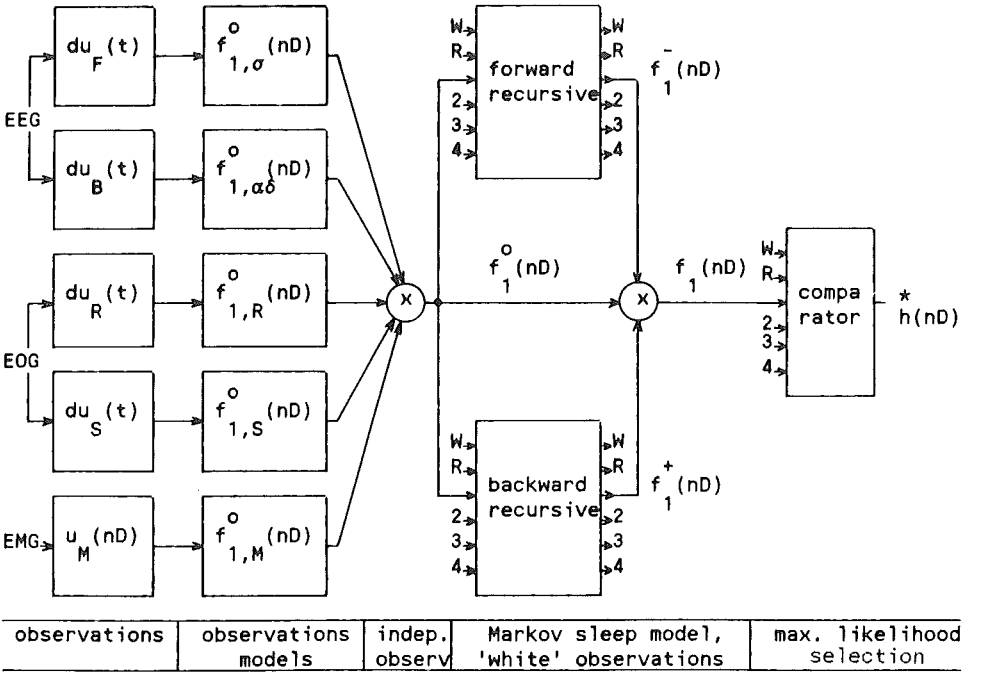


Figure 5: Block diagram of sleep stage monitor.
From left to right: vector of 'present' observations; computation of 'present' scalar observations likelihoods of sleep stage 1; multiplication of these yields the 'present' observations likelihood for sleep stage 1; smoothing with feedback of all (i.e. for all stages) thus computed 'present' likelihoods yields all 'past' and 'future' likelihoods; multiplication of these with the 'present' likelihoods provides all full observations likelihoods (shown for stage 1 only); the optimal stage decision is the one that corresponds to the largest likelihood, $f_j(nD)$.
The lower part of the figure shows the main propositions underlying each step.

The algorithm behaves as follows (see also figure 5). During a particular sleep stage, $h(nD)=j$, the observations models, (5)-(9), tend to generate observations, $U_X(nD,nD+D)$, that fit the corresponding scalar observations probability densities, $f^o_{j,X}(nD)$. For such observations, $f^o_{i,X}(nD)$ is maximal for $i=j$ $(i,j \epsilon S)$. According to (17) and (14), the same holds for the 'present' likelihood, $f^o_j(nD)$, and the 'full' likelihood, $f_j(nD)$.However, (14) shows that this tendency may be overruled if 'past' or 'future' likelihoods,

$f_i^-(nD)$ or $f_i^+(nD)$, clearly indicate another sleep stage, $i \neq j$. According to (15) and (16) this can only be the case for small transition probabilities, A, and 'past' or 'future' observations that strongly favour stage i. The recursive equations (15) and (16) also show that observations further away from the decision time, nD, are less powerfull in maximizing any particular 'past' or 'future' likelihood, $f_i^-(nD)$ or $f_i^+(nD)$, at time, nD.

V.4. Implementation, artifact rejection and parameter learning

The algorithms (14)-(22) that compute the likelihoods $f_i(nD)$, from the observations, U(0,ND), have been implemented on a minicomputer. The continuous time integrals in (18) and (19) were approximated in discrete time with an interval, Δ=0.01s, as follows:

$$\int_{t=nD}^{nD+D} z_{i,.}(t) \cdot du_{.}(t) \simeq \sum_{k\Delta=nD}^{nD+D-\Delta} \{z_{i,.}(k\Delta) \cdot \int_{t=k\Delta}^{k\Delta+\Delta} du_{.}(t)\} =$$

$$= \sum_{k\Delta=nD}^{nD+D-\Delta} \{z_{i,.}(k\Delta) \cdot [u_{.}(k\Delta+\Delta) - u_{.}(k\Delta)]\} =$$

$$= \sum_{k\Delta=nD}^{nD+D-\Delta} \{z_{i,.}(k\Delta) \cdot \Delta u_{.}(k\Delta)\} \qquad (23)$$

$$\int_{t=nD}^{nD+D} z_{i,.}^2(t) \cdot dt \simeq \sum_{k\Delta=nD}^{nD+D-\Delta} z_{i,.}^2(k\Delta) \cdot \Delta \qquad (24)$$

The forward increments, $\Delta u_{.}(k\Delta)$, can be reconstructed (based on the model of fig.2) from the EEG, e(t), either by an analogue circuit, or (as we did) by digital approximation as follows:

$$\Delta u_{.}(k\Delta) \simeq e(k\Delta) \cdot \Delta + \{e(k\Delta+\Delta) - e(k\Delta)\}/2\pi f_c \qquad (25)$$

The sampling interval, Δ, should not be much larger than 0.01s, because the frequency band of $s_{\cdot}(t)$, in particular $s_{\sigma}(t)$, may extend to about 20Hz. Nor should it be much smaller, because of the increasing risk of violating the white noise assumption in the model, which will affect the independence between $s_{\cdot}(k\Delta)$ and the increments, $\Delta\omega(k\Delta)=\Delta u_{\cdot}(k\Delta)-z_{h,\cdot}(k\Delta)\cdot\Delta$, of the Wiener processes, and thus cause bias (Kemp 1983[b]). For the same reason, no anti-aliasing filters should be applied.

For the final tuning of the monitor, we obtained sleep recordings (EEG, EOG and EMG) of one healthy female and one male patient, suffering from Huntington's Chorea, both aged 45. Using standard techniques, the recordings were made in the morning, following a night of sleep deprivation, and lasted 2½h. No prior knowledge of the subjects' sleep or sleep-related signals was available. No prior epoch selection or artifact rejection was applied. Sleep stage classification was performed (fig.6) with a time resolution, D=30s, both by 6 human specialists according to standard classification rules (Rechtschaffen and Kales 1968) and by the monitor.

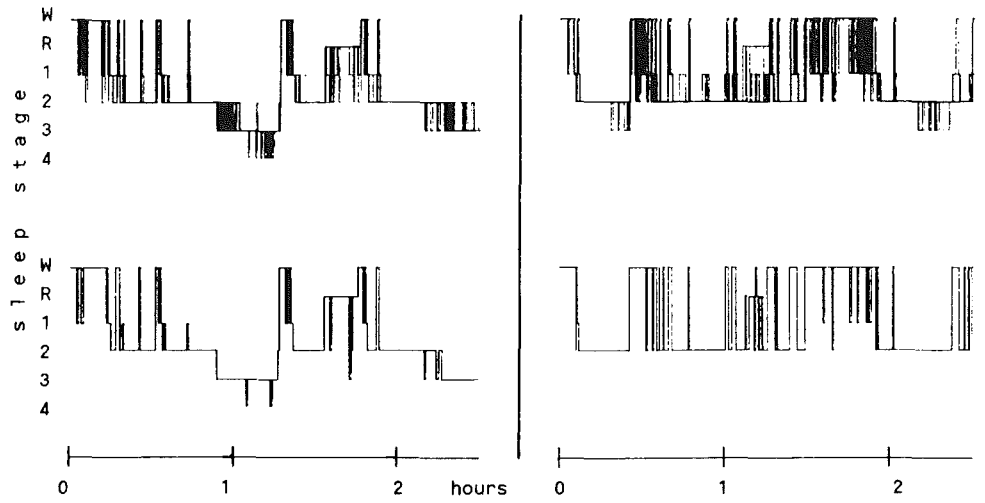


Figure 6: Sleep stage classifications of the learning set. Left: healthy female. Right: male patient. Upper: six human classifications superimposed. Lower: automatic classification.

The epochs of 30s in which all specialists disagreed with the monitor were considered to be misclassifications by the monitor. Many of these were caused by various EEG artifacts (i.e. differences between the real EEGs and the corresponding models). A nice way to cope with this would be to extend the models in such a way that they take artifacts into account. However, the resulting complexity would considerably reduce the interpretability of the models and the classification algorithms. We, therefore, adopted a simpler, more pragmatic approach. Because many artifacts do not last much longer than 1s, we first factorized the 'present' EEG likelihoods (18) and (19) into 1s segments as follows:

$$f^{o}_{i,F}(nD) = \prod_{md=nD}^{nD+D-d} f^{o}_{i,\sigma}(md) \qquad \text{,in which:} \qquad (26)$$

$$f^{o}_{i,\sigma}(md) = \exp\{ \int_{t=md}^{md+d} g_{\sigma}(i)\cdot s_{\sigma}(t)\cdot du_{F}(t) - \tfrac{1}{2} \int_{t=md}^{md+d} g^{2}_{\sigma}(i)\cdot s^{2}_{\sigma}(t)\cdot dt \} \qquad (26^{a})$$

and:

$$f^{o}_{i,B}(nD) = \prod_{md=nD}^{nD+D-d} [f^{o}_{i,\alpha}(md)\cdot f^{o}_{i,\delta}(md)\cdot f^{o}_{i,\alpha\delta}(md)] \qquad \text{,in which:} \qquad (27)$$

$$f^{o}_{i,\alpha}(md) = \exp\{ \int_{t=md}^{md+d} g_{\alpha}(i)\cdot s_{\alpha}(t)\cdot du_{B}(t) - \tfrac{1}{2} \int_{t=md}^{md+d} g^{2}_{\alpha}(i)\cdot s^{2}_{\alpha}(t)\cdot dt \} \qquad (27^{a})$$

$$f^{o}_{i,\delta}(md) = \exp\{ \int_{t=md}^{md+d} g_{\delta}(i)\cdot s_{\delta}(t)\cdot du_{B}(t) - \tfrac{1}{2} \int_{t=md}^{md+d} g^{2}_{\delta}(i)\cdot s^{2}_{\delta}(t)\cdot dt \} \qquad (27^{b})$$

$$f^{o}_{i,\alpha\delta}(md) = \exp\{ \int_{t=md}^{md+d} -g_{\alpha}(i)\cdot g_{\delta}(i)\cdot s_{\alpha}(t)\cdot s_{\delta}(t)\cdot dt \} \qquad (27^{c})$$

Although clear and moderate alpha rhythms are typical for stage W and for stages R and 1, respectively, they are sometimes completely inhibited by altertness during stage W. Also many subjects show no alpha rhythm at all in stage W, but do occasionally in stages 1 and R. Alpha-like activity may even be present in stages 3 and 4 (Hauri and Hawkins 1973, Scheuler et al 1983).

Delta waves show similar inconsistencies. Although they are typical for stage 2, delta waves are generally not present during the early parts of stage 2. Because of these inconsistencies, alpha and delta waves caused considerable misclassifications via their contributions to (19), i.e. (27). The misclassifications caused by alpha-like activity combined with delta waves in stages 3 and 4 could be avoided by setting $f^{o}_{i,\alpha\delta}(md)=1$. The misclassifications (into stage 2) caused by occasionally low alpha powers during stages W, R and/or 1, were avoided by setting a lower limit to the corresponding alpha contribution as follows:

$$f^{o}_{j,\alpha}(md) = \max[f^{o}_{j,\alpha}(md), f^{o}_{2,\alpha}(md)] \qquad j\epsilon\{W,R,1\} \qquad (28)$$

The misclassifications (into stage W) caused by occasional absence of low-frequency activity in stage 2 were avoided by a similar low limit for the corresponding delta contribution:

$$f^{o}_{2,\delta}(md) = \max[f^{o}_{2,\delta}(md), f^{o}_{W,\delta}(md)] \qquad (29)$$

The contributions of the sigma rhythm to the exponents of (18) and (26) can be driven to excessively large negative values due to high-frequency muscle activity in the vicinity of the frontal EEG electrode. The resulting misclassifications can be avoided by setting a low limit for the sigma contribution as follows:

$$f^{o}_{j,\sigma}(md) = \max[f^{o}_{j,\sigma}(md), C_{-3SD} \cdot f^{o}_{W,\sigma}(md)] \qquad j\epsilon\{2,3,4\} \qquad (30)$$

$$\text{in which} \quad C_{-3SD} = \exp\{-2\pi B \cdot d/4 - 3\sqrt{2\pi B \cdot d/2}\} \qquad (30^{a})$$

These artifacts are often accompanied by body movements that cause low frequency electrode artifacts and thus large positive delta contributions, $f^{o}_{i,\delta}(md)$. Therefore, during such intervals, i.e. when the low limit (30) was reached, we set the following upper limit for the delta contributions:

$$f^{o}_{j,\delta}(md) = \min[f^{o}_{j,\delta}(md), f^{o}_{W,\delta}(md)] \qquad j\epsilon\{2,3,4\} \qquad (31)$$

		human						automatic					
		W	R	1	2	3	4	W	R	1	2	3	4
human	W	13.8	0.0	1.9	0.7	0.0	0.0	15.5	0.0	0.4	0.4	0.0	0.0
	R	0.0	2.7	0.5	0.8	0.0	0.0	0.1	3.3	0.0	0.5	0.1	0.0
	1	1.9	0.5	7.1	3.3	0.0	0.0	6.6	0.6	2.9	2.6	0.1	0.0
	2	0.7	0.8	3.3	48.7	2.1	0.0	3.4	0.9	0.5	48.5	2.2	0.0
	3	0.0	0.0	0.0	2.1	8.4	0.4	0.0	0.0	0.0	1.0	9.6	0.3
	4	0.0	0.0	0.0	0.0	0.4	0.1	0.0	0.0	0.0	0.0	0.4	0.1

Figure 7: Comparability matrices of the learning set. An element in row i and column j denotes the percentage of 30s epochs classified as stage i by a human and as stage j by another human (left matrix) or by the automatic (right matrix) classifier. The human vs human comparabilities (left matrix) are estimated by averaging the percentages over all (30) possible pairs of human classifiers. The human vs automatic comparabilities (right matrix) are estimated by averaging over the (6) possible pairs of a human with the automatic classifier. Note the overclassification into Wakefulness by the automatic classifier.

The transition rate, $a=1/30s^{-1}$, and the parameters in the sleep characteristics matrix of table 1 have been found by trial and error. The goal was to maximize the average agreement, which is the average (over the six human classifiers) percentage of time (i.e. of the 2½ hours) during which the monitor and the classifier agree in their classification. The final average agreements were 84.0% and 75.7% for the healthy subject and the patient,

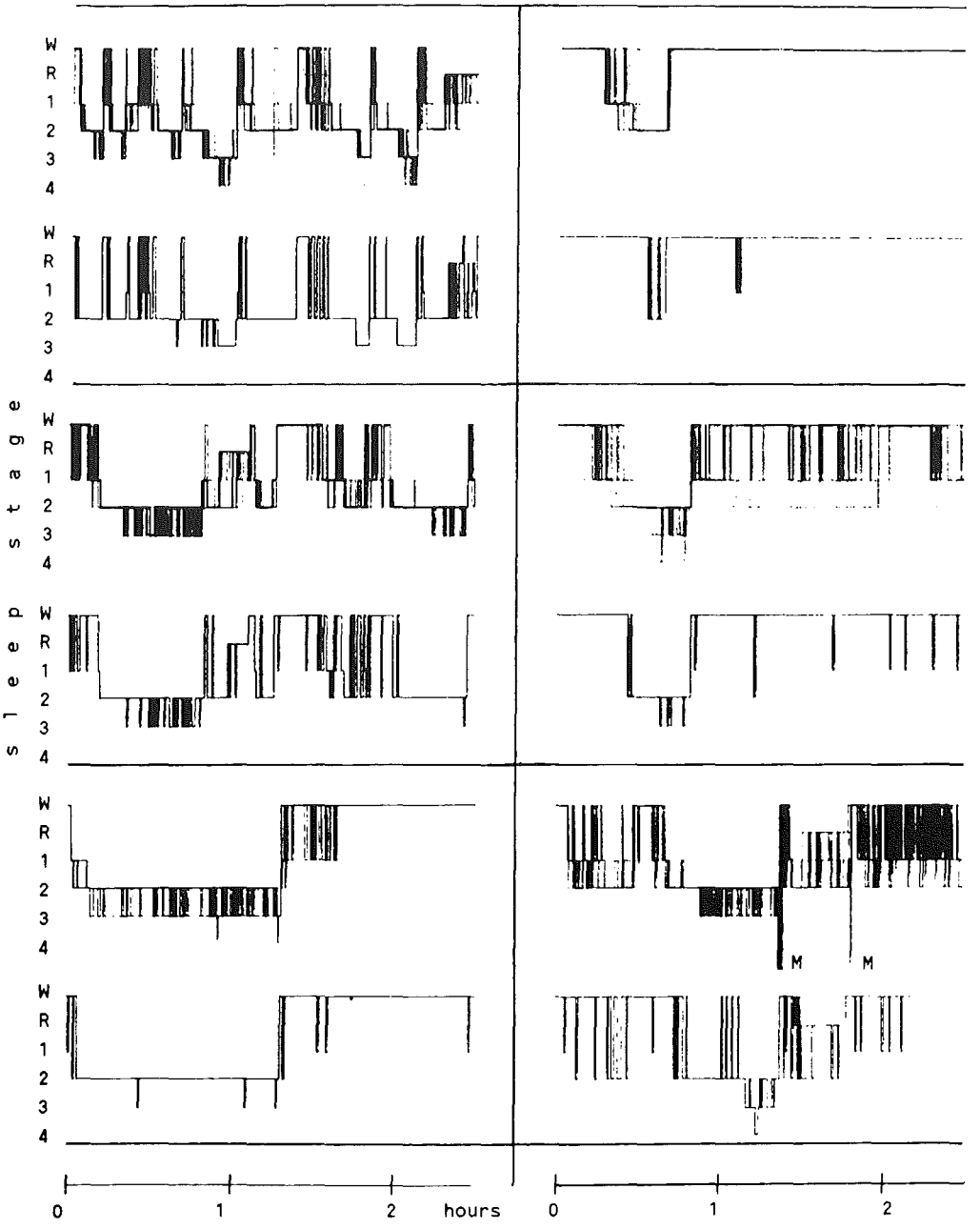


Figure 8: Human (upper traces within each box, superimposed) and automatic (lower traces) classifications of the 3 healthy volunteers (left) and 3 patients (right) in the test set. M: movement time.

respectively. The corresponding average (over the 30=5x6 possible pairs out of the 6 human classifiers) agreements between human classifiers were 82.8% and 78.8%, respectively. When averaged over the two subjects, the human-human and human-automatic agreement percentages were 80.8 and 79.9, respectively. These percentages are the sums of the diagonal elements of the comparability matrices of figure 7.

V.5. Performance

Sleep recordings, classifications (fig.8), and comparabilities (fig.9) were obtained from a test set by the same methods as applied to the learning set. The test set consisted of three healthy males (aged 32, 34 and 47) and two male and one female patient suffering from Huntington's Chorea (aged 28, 47 and 48). The average human-automatic agreements were 75.4% for the healthy subjects and 65.9% for the patients group. The corresponding human-human agreements were 77.3% and 72.9%, respectively. When averaged over the full test set, the human-automatic and human-human agreements were 70.6% and 75.1%, respectively.

		human						automatic					
		W	R	1	2	3	4	W	R	1	2	3	4
human	W	34.7	0.1	3.1	1.7	0.0	0.0	37.9	0.0	1.3	0.5	0.0	0.0
	R	0.1	1.8	0.4	0.6	0.0	0.0	0.2	1.6	0.2	0.9	0.0	0.0
	1	3.1	0.4	7.1	3.4	0.0	0.0	10.0	0.3	1.3	2.5	0.0	0.0
	2	1.7	0.6	3.4	27.6	2.9	0.0	6.8	0.8	0.2	27.1	1.4	0.1
	3	0.0	0.0	0.0	2.9	3.8	0.2	0.0	0.0	0.0	4.1	2.8	0.0
	4	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.2	0.0

Figure 9: Average (over the test set and the 30 possible pairs of human classifiers) human-human, and average (over the test set and all (6) possible pairs of humans vs the automatic classifier) human-automatic comparability matrices.

We also computed comparability matrices individually for each subject. The individual average agreement percentages, i.e. the sum of the diagonal elements of these individual comparability matrices, are depicted in table 2. Rather consistent (over the learning and test set) automatic misclassifications into stage Wakefulness originated from not using the information from theta activity (during stage 1) and the "alpha-drop" that sometimes accompanies the transition from wakefulness to stage 1. Because the difference between W and 1 is not too important from a diagnostic point of view, we also computed the agreement percentages when this difference was ignored (table 2). In this way, the average (over the full test set) human-automatic and human-human agreements reached 81.8% and 81.3%, respectively.

Although this comparability analysis shows how close the automatic and human classifications of the test set are, it does not show how close either of them is to real sleep. An indication of the performance that can be achieved if the sleep-related observations really behave like our models, is obtained as follows. We simulated the continuous-time Markov sleep process, h(t), with individual time-varying rates (Kemp and Kamphuisen 1986: fig.5c, stage M neglected). The resulting 'hypnograms' closely resemble real ones of a healthy young subject. We also simulated the observations models (5-9), and modulated them every second by h(t). The simulated observations were fed into the monitor and the resulting automatic sleep classification (resolution 30s) was compared, with a time resolution of 1s, to the simulated hypnogram. From ten 'nights' of 8 hours this resulted in the average comparability matrix of figure 10.

					W = 1		
		human	aut-hum	automatic	human	aut-hum	automatic
learning set	HS	82.8	1.2	84.0	84.2	2.0	86.2
	HC	78.8	-3.1	75.7	84.8	2.1	86.9
test set	HS	74.9	-5.2	69.7	78.9	-4.7	74.2
	HS	78.6	-3.7	74.9	82.0	1.9	83.9
	HS	78.4	3.1	81.5	81.2	4.6	85.8
	HC	95.7	-7.5	88.2	97.5	-3.1	94.4
	HC	52.6	9.3	61.9	72.0	10.9	82.9
	HC	70.3	-22.8	47.5	76.1	-6.3	69.8
	AV	75.1	-4.5	70.6	81.3	0.5	81.8

Table 2: Average agreement (%) between humans (first column) and between human and automatic (third column) sleep classification for each subject of the learning set and test set. The fourth and sixth columns show the agreements if Wakefulness and stage 1 are taken to be the same stage. The second and fifth column show the differences between automatic vs human and human vs human agreements. HC: patient, HS: healthy subject, AV: average over the test set.

		classification					
		W	R	1	2	3	4
simulation	W	3.7	0.0	0.3	0.0	0.0	0.0
	R	0.4	21.2	0.2	0.0	0.0	0.0
	1	3.1	0.1	5.1	0.2	0.0	0.0
	2	1.3	0.1	0.5	48.2	0.1	0.0
	3	0.0	0.0	0.0	4.6	4.3	0.2
	4	0.0	0.0	0.0	1.8	2.4	2.1

Figure 10: Average (over ten simulated 'nights') comparability matrix of simulated versus automatically classified sleep stages. Agreement 84.6%

V.6. Discussion

The optimal monitor consists of preprocessors for each sleep-related signal component (e.g. sigma-rhythm, REMs etc.). These preprocessors each compute a sequence of corresponding probability densities. They are followed by smoothing operations with feedback, (14)-(17), that combine these densities in such a way that the clearest information is weighed heaviest. For instance large, clear sigma spindles dominate an accidentally occurring rapid eye movement, resulting in a stage 2, 3 or 4 decision. Also clear information from 'past' or 'future' dominates doubtful 'present' information. This procedure mimics the human procedure and strongly contrasts with the commonly applied technique of taking intermediate decisions on sigma-presence etc. The latter does not account for the reliability of the various subdecisions. The smoother operates similarly to the 'BAMPS-algorithm' (Haralick 1983), which however does not account for the causal feedback in the EEG model. A slightly different algorithm, similar to the Viterbi-algorithm (Haralick 1983), has been applied recently to EEG only (Lacroix and Stanus 1985[a,b], Lacroix 1985), but it does not properly account for 'future'observations.

Only a few automatic methods have been compared to the standardized (Rechtschaffen and Kales 1968) human classification (Martin et al 1972, Lacroix and Stanus 1985[a,b], Hasan 1983, Hasan 1985 (review), Johnson 1977), resulting in human-automatic agreements ranging from 65% to 86% for normal and from 63% to 75% for disturbed sleep. The average agreements for our healthy subjects and patients (as derived from the test set in table 2), i.e. 75.4% and 65.9%, are within these ranges. However, sleep disturbances, artifacts and 'difficult' or 'easy' sleep-related signals (e.g. with unclear or clear sigma rhythm) will considerably influence both the human-automatic and the human-human agreement. In order to decide whether the automatic method is a feasible alternative to the human procedure, it is therefore more

appropriate to compare these two agreements to each other. So far, this has only been done in two studies (Martin et al 1972, Lacroix and Stanus 1985[a,b]), resulting in human-human agreements that were in both cases 8% better than the corresponding human-automatic agreements of 81% and 86%, respectively. In our group, this difference is only 4.5%.

The monitor performs better on simulated than on real stage 1. The difficulty in separating stages W and 1 is probably due to poor modelling and may be partly caused by unrealistic assumptions (e.g. about the difference in alpha activity and slow eye movements between these two stages, or about the duration of sleep stages being a multiple of 30s) already underlying the standard classification rules (Herrmann and Kubicki 1984, Hasan 1983, Molinari et al 1984). This difficulty is shared with many other systems (Johnson 1977, Herrmann and Kubicki 1984 (overviews)). If we ignore the difference between stages W and 1, which for diagnostic purposes, is fairly unimportant, our human-automatic agreement rises to 81.8% (table 2), which is slightly (0.5%) better than the corresponding human-human agreement.

Figure 9 shows that the most prominent difference between human classification and the monitor is due to automatic classification into Wakefulness of epochs that are considered as stage 1 or 2 by specialists. Inspection of these epochs showed, that these are partly characterized by theta activity and K-complexes, features that were not yet included in the model. Therefore, it may be useful to extend the monitor with likelihoods that are based on models of these features.

The basic advantages of the weighing algorithm, the fact that the first practical results already seem better than those described for other methods, and the possibility of further improvements suggest that the model-based approach of developing a sleep stage monitor is sound. However, evaluation on a larger test set, whole-night recordings and dif-

ferent age- and patient-groups must be performed before we can rely on the method in clinical studies.

Disagreement between computer and specialist does not imply that the automatic procedure is wrong; in particular the human capability of recognizing sigma rhythm seems worse than that of some automatic methods (Gaillard and Blois 1981, Dumermuth et al 1972). A closer relationship to 'real', physiological sleep may be obtained via more physiological models or better time resolution D. The theoretical framework in this paper allows any time resolution and is of sufficient generality to include more or different observations models. However, a strong limitation remains the doubtful, although generally applied, assumption of synchronous discontinuous transitions of all sleep characteristics (Rechtschaffen and Kales 1968, Molinari et al 1984, Parmeggiani et al 1985, Gath and Bar-On 1983).

Acknowledgement. We gratefully acknowledge the support and cooperation of Professor Dr. F.H. Lopes Da Silva, Professor Dr. H.A.C. Kamphuisen, Dr. E.L.E.M. Bollen, Mrs. H.D.J. Geurts, Mrs. G.J.M. Keulen, Mrs. M. Van Rijnsbergen and Mrs. H.M.T. Kemp.

VI.MISCELLANEOUS RESULTS

This chapter briefly describes some parallel results that may find application in related fields. It should be read in connection with the preceding chapters. The references are not the result of a thorough literature search.

VI.1. Estimation of the feedback gain of EEG rhythm generators

Based on the model of figure 1 in chapter III, we have a discretized observations equation (Δ=0.01s):

$$\Delta u(t) = p(t) \cdot s(t) \cdot \Delta + \Delta\omega(t) \qquad (1)$$

where the 'forward' Wiener process increments, $\Delta\omega(t)=\omega(t+\Delta)-\omega(t)$ are independent of the feedback signal, s(t). We assume that the time-dependent feedback gain, p(t), shows the same time course in a number, n, of mutually independent experiments, indexed i.

Then the maximum likelihood estimator of the feedback gain reads:

$$\hat{p}(t) = \sum_{i=1}^{n} \{s_i(t) \cdot \Delta u_i(t)\} / \sum_{i=1}^{n} \{s_i^2(t) \cdot \Delta\} \qquad (2)$$

This method has been applied to estimate, from n experiments, the flash-evoked attenuation of the feedback gain of the alpha rhythm generator in healthy volunteers (Kemp 1983[b]). In this way, the latency from the flash to the first sign of attenuation could be computed with a time resolution of 5ms. All other methods reported in the literature had resolutions >50ms. At that time, we applied 'backward' increments, $\breve{\Delta}\omega(t)=\omega(t)-\omega(t-\Delta)$, that were dependent on s(t). This results in considerable bias in equation (2). For instance, simulations with a gain, p(t)=0, resulted in an estimate, $\hat{p}(t)=1$.

(The bias correction heavily relied on the correctness of model assumptions.) New simulations (fig.1) and also practical applications, now using forward increments, $\Delta u(t)$, show a considerably reduced bias. Anti-aliasing filters must not be applied since these cause bias (fig.1) for the same reason.

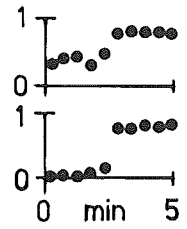


Figure 1: Estimate (vertical axis), $\hat{p}(t)$, of α-feedback gain from simulated EEG observations. The simulations are made using the model of figure 1 in chapter III, with a feedback gain, $p(t)$, that switches after 2.5 minutes from 0.00 to 0.85. Upper and lower trace are with and without anti-aliasing filtering (50Hz, 24dB/oct), respectively.

VI.2. Dissociation and continuous variation of sleep characteristics

Based on the likelihoods, $f^{o}_{i,\rho}(md)$ with $\rho\varepsilon\{\sigma,\alpha,\delta,M\}$ and $i\varepsilon S$, as defined in chapter V (formulas 26^{a}, $27^{a,b}$ and 22), we define sigma-, alpha-, delta-, and EMG-indicators by their log-likelihood ratios, $\ln\{f^{o}_{k,\rho}(md)/f^{o}_{1,\rho}(md)\}$, with time resolution, D=d=1s. The feedback gains, $g_{\cdot}(k)$ and $g_{\cdot}(1)$, and the muscle twitch probabilities, $q_{M}(k)$ and $q_{M}(1)$, are the highest and the lowest values from table 1 of chapter V, respectively. Each rapid and slow eye movement (REM and SEM) detected is indicated by a vertical bar.

Figure 2 shows that low muscle tone is reached (i.e. the EMG-indicator drops to its lowest level) a few minutes before stage REM is entered. The δ-indicator reflects the well-known gradual increase in δ-power which was seen in all subjects who reached stage 3. These two examples illustrate the desynchronization and non-jumping behaviour of some sleep-related variables as was already discussed in chapters I and V. The log-likelihood ratios offer a tool to study these in conjunction with automatic sleep classification.

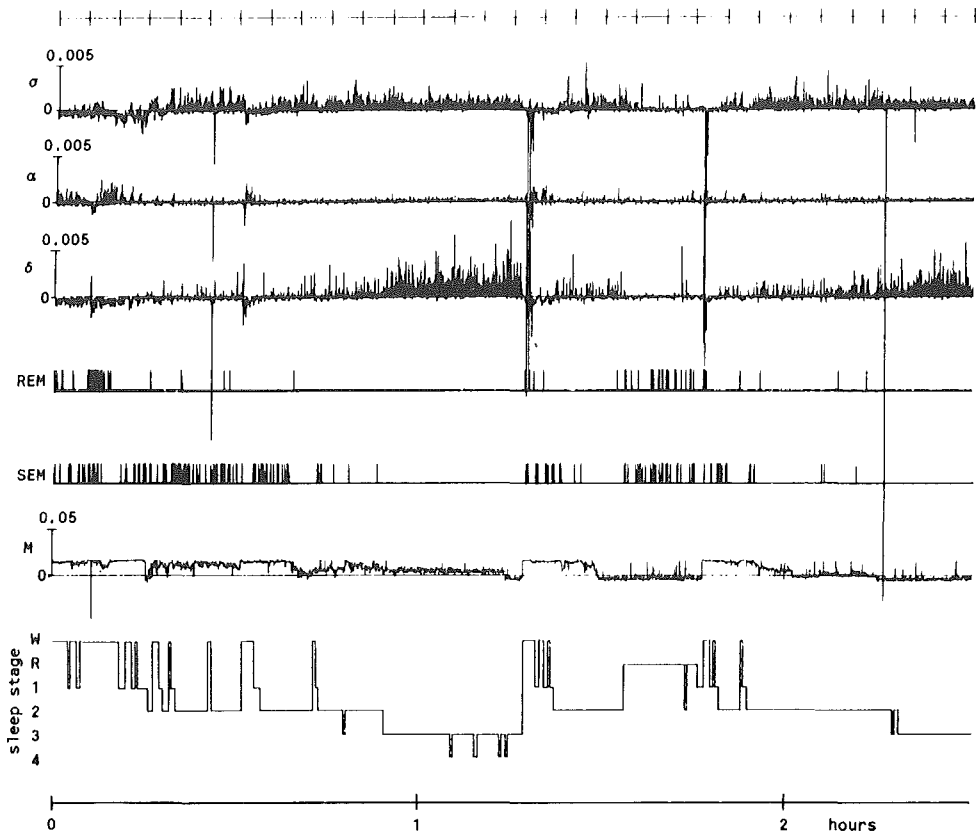


Figure 2: From upper to lower trace: sigma-, alpha-, delta-, REM-, SEM- and EMG-indicator and human consensus sleep classification for a subject of the learning set.

VI.3. Alpha-delta sleep

Power spectral analyses (fig.3) of two EEG signals demonstrated the presence of an alpha-like component (frequency about 10Hz), mainly at fronto-central, but also spreading to occipital regions of the head, during sleep stages 3 and 4 in some subjects. This confirms the results of Hauri and Hawkins (1973) and of Scheuler et al (1983). Apparently, we can only be sure of the absence of alpha-like activity during sleep stage 2.

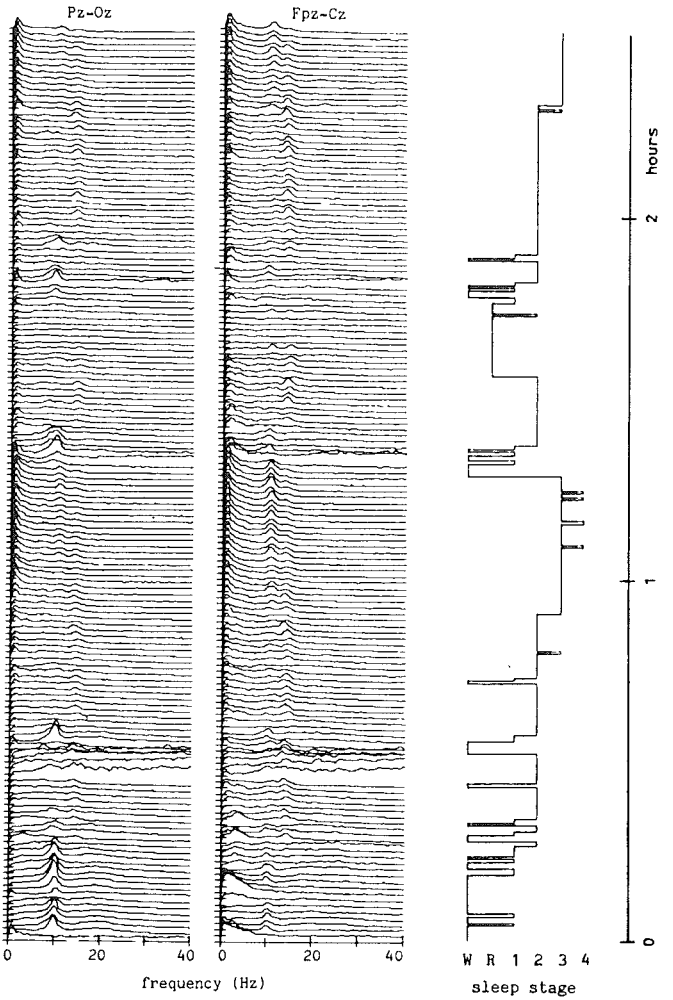


Figure 3: Sequential (one per minute) power spectral analyses of pariëto-occipital (Pz-Oz, left column) and fronto-central (Fpz-Cz, middle column) regions of the scalp of the same subject as in figure 2. Frequency(f)-dependent √power calibration: $20/\sqrt{\{f \cdot (f+10)\}}\,\mu V/\sqrt{Hz}$ per vertical division. Right column: human consensus sleep classification. Note the pariëto-occipital 10Hz alpha rhythm during Wakefulness and the fronto-central 14Hz sigma rhythm during stages 2, 3 and 4. Note also the fronto-central 10Hz rhythm during stages 3 and 4.

VI.4. Ambulatory cassette recordings

As was already put forward in chapter I, major costs in sleep research are due to the hospitalization for the purpose of recording the EEG etc. during the night. In addition to obvious general problems, such hospitalization also influences the sleeping system, the sleep disorder and hence the diagnosis.

It has recently become possible to acquire the required signals at the patients home, unsupervised, and transmit them by radio telemetry (license required) and/or telephone line (expensive) to the laboratory. Another possibility is by direct recording on a 'Walkman-like' cassette recorder which the patient carries on a belt (Stefan and Burr 1982, Campbell and Wilkinson 1981, Riley and Peterson 1983, Sewitch and Kupfer 1985, De Groen et al 1985, Koffler and Gotman 1985, Jayakar et al 1985). For reasons of time and costs we chose the latter appróach, based upon a commercially available system that was already present in our laboratory. It can record 25 hours of 4 analogue signals, each with a bandwidth that we extended to 0.5-100Hz and a signal-to-noise ratio of 35dB on a standard audio cassette.

The two EEG signals and one horizontal EOG signal that are required for sleep stage monitoring are recorded directly. We built in a multiplexer with 4 input channels that drives the remaining recorder channel (fig.4). The input channels are each sampled with 1Hz and are used for event marker, body temperature (accuracy 0.1°C), respiration (air flow), and muscle tone. The latter is obtained by smoothing the rectified high-frequency component of the EMG. In this way the limited bandwidth of the cassette recorder (which is even more limited in the commercially available 8-channel recorder) poses no problem. The multiplexer is crystal-controlled; therefore, triggering the AD-converter by the demultiplexer results in accurate ($\pm$ 15s per 24h) real-time analysis, independent of tape speed.

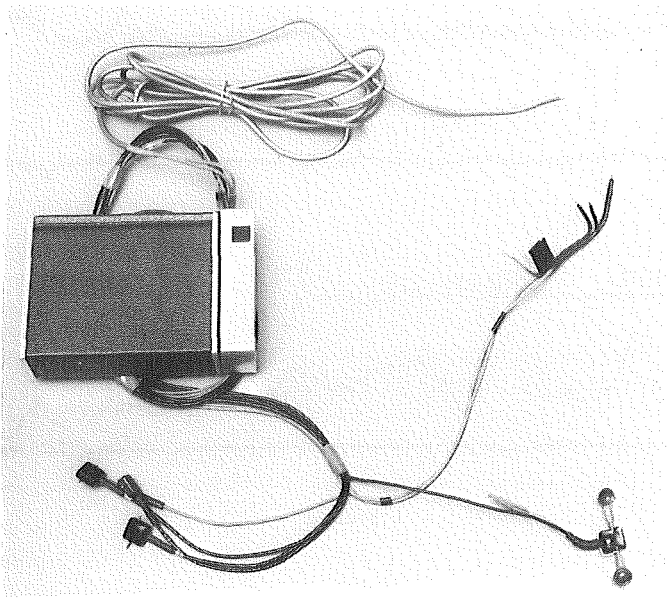


Figure 4: Ambulatory cassette recorder with built-in multiplexer and event marker. Connected are ground wire and head-on preamplifiers (lower left) for EEG (2x) and EOG, rectal body temperature probe (upper right), thermistorcombination for oral-nasal respiration air flow (right) and head-on amplifier with electrodes for the submental EMG (lower right). Not shown are the 1 ground-, 6 EEG- and 2 EOG-electrodes and also the belt by which the recorder can be carried.

Five recorders are now available in our department and in the Department of Physiology. Since 1985, about 100 nights (e.g. fig.5) have been recorded at the patients homes. Because the recordings take place in ones own, familiar environment, the recorded sleep is more natural. A first (extra) adaptation night, which is generally required if the registrations are obtained in the hospital, is therefore not necessary. In some cases, the application of electrodes and other transducers is also performed at the patients home. This gives us the opportunity to observe some sleep-related environmental factors. The recordings do not suffer from more artifacts than clinical recordings do. They are of sufficient quality to allow visual sleep stage classification.

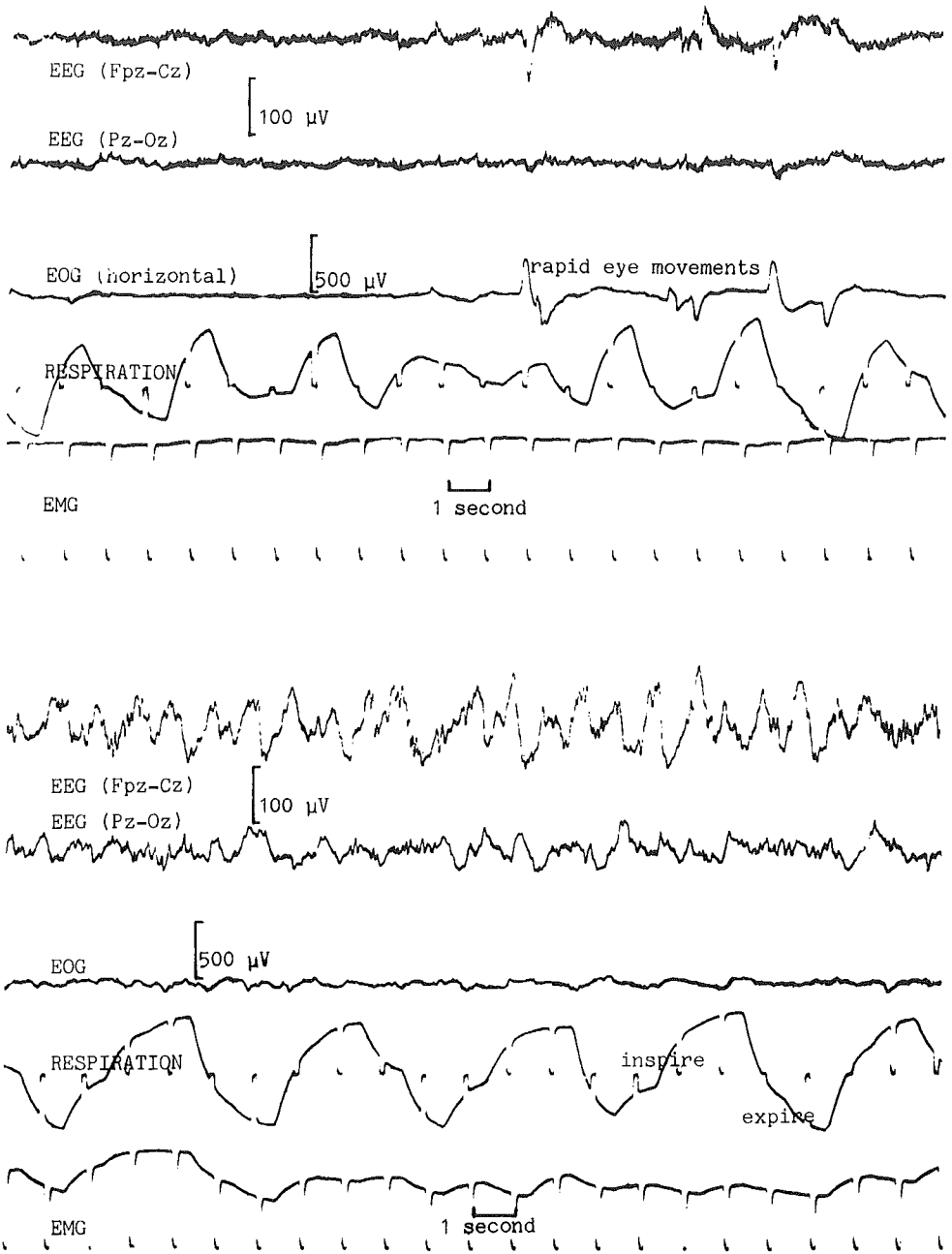


Figure 5: Cassette recordings during wakefulness (upper half; note the high muscle tone and the rapid eye movements) and during sleep stage 4 (lower half; note the smaller muscle tone and the lowfrequency activity in the EEG).

Replaying the cassettes into a computer and visual classification on a high-resolution display, enables one to subject the hypnograms directly to statistical analysis. Using this approach, a considerable amount of working hours, both for supervision and for data handling, can be saved.

VI.5. Continuous-time monitoring

Although there is no theoretical limit to the time resolution, D, of the discrete time monitor of chapter V, it will be practically limited by the maximum achievable computation speed. In very fast environments one might therefore be interested in the structure of the continuous-time monitor. We will restrict this section to the causal part of the monitor that takes decisions, based on past observations only. Although there is no theoretical necessity for this restriction, it serves the simplicity of reasoning and allows implementation by analogue circuits.

In chapters III and IV we derived a differential equation that yields the log-likelihood ratio for the optimal detection of the state of a binary Markov chain, $p(t)\epsilon\{0,1\}$. Because the a posteriori probabilities:

$$P_i^-(t) = P\{p(t)=i \mid \text{observations up to time } t\} \qquad i\epsilon\{0,1\} \qquad (3)$$

were functions of the 'past-observations' likelihoods:

$$f_i^-(t) = f\{\text{observations up to time } t \mid p(t)=i\} \qquad i\epsilon\{0,1\} \qquad (4)$$

this differential equation can simply be transformed into one that yields the test statistic:

$$R_{10}(t) = \ln\{P_1^-(t)/P_0^-(t)\} \qquad (5)$$

$R_{10}(t)$ directly indicates which one of the two a posteriori probabilities is the larger. The differential equation for $R_{10}(t)$ describes an integrator with exponential feedback and is driven by a function, $dc_{10}(t)$, of the 'present observations' only. It reads:

$$dR_{10}(t) = [\{a_{1|0} \cdot e^{-R_{10}(t)} - a_{0|1}\} - \{a_{0|1} \cdot e^{R_{10}(t)} - a_{1|0}\}] \cdot dt + dc_{10}(t) \qquad (6)$$

where $a_{j|i}$ are the transition rates of the Markov chain.

In another paper (Kemp and Jaspers 1984) we derived the set of differential equations that yield the test statistics, $R_{ji}(t)=\ln\{\bar{P}_j(t)/\bar{P}_i(t)\}$, for the optimal detection of the state of an M-ary Markov chain, $p(t)\epsilon S$. These differential equations are also integrators with exponential feedback, driven by direct memoryless functions, $dc_{ji}(t)$, of the observations. They read:

$$dR_{ji}(t) = [\sum_{k\neq j} \{a_{j|k} \cdot e^{-R_{jk}(t)} - a_{k|j}\} - \sum_{k\neq i} \{a_{i|k} \cdot e^{R_{ki}(t)} - a_{k|i}\}] \cdot dt + dc_{ji}(t)$$
$$i,j,k\epsilon S \qquad i\neq j \qquad (7)$$

The drivers, $dc_{ji}(t)$, are the log-likelihood ratios:

$$dc_{ji}(t) = \ln\{f^o_j(t)/f^o_i(t)\} \qquad (8)$$

with 'present-observation' likelihoods:

$$f^o_i(t) = f\{\text{observation at time } t|p(t)=i\} \qquad i\epsilon S \qquad (9)$$

in which the observation may be a vector observation. The simultaneous computation of M-1 ratios, e.g. all $R_{j1}(t)$ for one fixed $l\epsilon S$, is sufficient to

decide which a posteriori probability, $P_i^-(t)$, is largest, i.e. in which state ,$i \epsilon S$, the Markov chain most probably resides at time t. The other $R_{ji}(t)$ for $i \neq 1$, that are required for the feedback, can be computed from the $R_{j1}(t)$ as follows (from the definition of $R_{ji}(t)$):

$$R_{ji}(t) = R_{j1}(t)+R_{1i}(t) = R_{j1}(t)-R_{i1}(t) \qquad i,j \epsilon S \qquad (10)$$

Although this M-ary detector has been derived for white noise corrupted observations of a priori known signals only, chapter V shows that the discrete-time (time interval D) smoother (i.e. the integrator with feedback) is not influenced if Poisson point processes and causal feedback of the observations enter the models. The M-ary detector can probably also be derived by taking the limit ($D \to 0$) of the discrete time monitor, as was done earlier for the binary case (Kemp and Blom 1981).

VI.6. Calibration of standard white noise

Because the monitor is nonlinear, the power of the drivers is important. The 'cortical EEG' signals, du(t), or $\Delta u(t)$ in discrete time, should be the increments of a standard Wiener process when EEG rhythms are absent, as was assumed in the derivation of the algorithms.

Analogue, electric, standard white noise, i.e. the derivative of the standard Wiener process, should have a flat two-sided power spectral density, PSD=$1V^2$/Hz (if we take volts and seconds as standard dimensions). Driving the circuit of figure 6 by non-standard white noise, du(t)/dt, yields a signal:

$$DC \simeq E\{h(t)\} = E\{g^2(t)\} = E\{ \int_{\tau=0}^{\infty} 2\pi f_k \cdot \exp(-2\pi f_k \cdot \tau) \cdot du(t-\tau) \}^2 =$$

$$= \int_{\tau=0}^{\infty} \{2\pi f_k \cdot \exp(-2\pi f_k \cdot \tau)\}^2 \cdot E\{du(t-\tau)\}^2 =$$

$$= \int_{\tau=0}^{\infty} \{2\pi f_k \cdot \exp(-2\pi f_k \cdot \tau)\}^2 \cdot PSD \cdot d\tau =$$

$$= \pi \cdot f_k \cdot PSD \qquad (11)$$

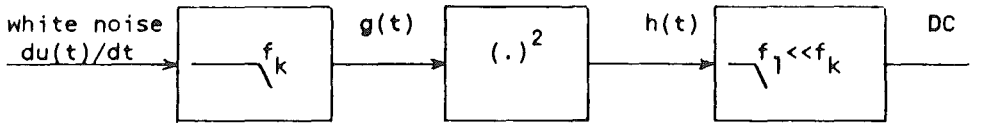


Figure 6: Measurement of white noise power spectral density. Both causal lowpass filters are first order. The first one has a cut-off frequency, f_k, and impulse response function, $\gamma(\tau)=2\pi f_k \cdot \exp(-2\pi f_k \cdot \tau)$. The noise is white up to a much higher frequency than f_k.

In practice the dynamic range of electronic circuits forces us to calibrate the PSD to values much smaller than $1V^2/Hz$. The required rescaling to correct this is partly performed by the AD-converter where we also get rid of the volt as arbitrarily chosen reference dimension. We applied this method to calibrate the EEG signals, to stabilize (by feedback) the analogue white noise generator that was applied in the simulations of chapter V, and to calibrate our power spectral analysis system.

Digital, discrete time with interval Δ, increments of the non-standard Wiener process, u(t), should obey:

$$\frac{1}{n} \cdot \sum_{t=\Delta}^{n\Delta} \{\Delta u(t)\}^2 \simeq E\{\Delta u(t)\}^2 = \Delta \cdot PSD \qquad (\Delta \text{ in seconds}) \qquad (12)$$

This equation was applied to check the above mentioned calibrations.

N.B. Equation (12) no longer holds if presampling anti-aliasing filtering is applied.

VI.7. Analogue implementation of stochastic differential equations

Several results in stochastic filtering theory are in the form of a differential equation like:

$$d\hat{\beta}(t) = a\{\hat{\beta}(t)\}\cdot dt + b\{\hat{\beta}(t)\}\cdot d\omega(t) \tag{13}$$

where

$$d\hat{\beta}(t) = \lim_{\delta\to 0}\{\hat{\beta}(t+\delta)-\hat{\beta}(t)\} \tag{13^a}$$

$$d\omega(t) = \lim_{\delta\to 0}\{\omega(t+\delta)-\omega(t)\} \tag{13^b}$$

$\omega(t)$ is a standard Wiener process (13^c)

$a\{.\}$ and $b\{.\}$ are 'nice' functions (13^d)

According to Larson and Shubert (1979:th.4.5.1), equation (11) in chapter III (with $du(\nu)-\hat{\beta}^{-}(\nu)s(\nu)d\nu=d\omega(t)$) is such a differential equation. The forward (with respect to $\hat{\beta}(t)$) differentials, $d\omega(t)$, are often difficult to realize (as in chapter III) by analogue (causal) systems. They are essential because only forward differentials are independent of $\hat{\beta}(t)$, implying that $E[b\{\hat{\beta}(t)\}\cdot d\omega(t)]=0$.

However, according to Itô's differentiation rule (Larson and Shubert 1979: ch.6.5), a function, $\gamma\{\hat{\beta}(t)\}$ of $\hat{\beta}(t)$, having the derivative:

$$\delta\gamma\{\hat{\beta}(t)\}/\delta\hat{\beta}(t) = b^{-1}\{\hat{\beta}(t)\} \tag{14}$$

obeys the differential equation:

$$d\gamma\{\hat{p}(t)\} = [\delta\gamma\{\hat{p}(t)\}/\delta\hat{p}(t)]\cdot d\hat{p}(t) + \tfrac{1}{2}[\delta^2\gamma\{\hat{p}(t)\}/\delta\hat{p}^2(t)]\cdot b^2\{\hat{p}(t)\}\cdot dt =$$

$$= [b^{-1}\{\hat{p}(t)\}\cdot a\{\hat{p}(t)\} - \tfrac{1}{2}[\delta b\{\hat{p}(t)\}/\delta\hat{p}(t)]]\cdot dt + d\omega(t) \qquad (15)$$

in which the forward differentials, $d\omega(t)$, are no longer multiplied by a function of $\hat{p}(t)$, and therefore (15) can be interpreted as an ordinary, physically implementable, white noise driven differential equation (Wong and Zakai 1965, Jazwinsky 1970:ch.4).

Formula (14) is a recipe that yields a transformation, $\gamma\{\hat{p}(t)\}$ of $\hat{p}(t)$, that can be obtained via an analogue implementation of a differential equation. In chapter III this recipe resulted, except for a constant and the supposedly a priori known signal $s(\nu)$, in the log-likelihood ratio, $\gamma\{\hat{p}(t)\}=\ln\{\hat{p}(t)/(1-\hat{p}(t))\}$. The differential equation (15) for its computation was simpler and easier to interpret than direct computation of $\hat{p}(t)$; compare equations (16) and (11) in chapter III.

VI.8. Some details of our particular implementation

The automatic sleep stage classification procedure of chapter V may be realized in various ways. Our particular implementation was based on available experience and material, like an analogue tape recorder, an LSI11/23 minicomputer system and electronic preprocessing circuits.

VI.8.1. Tape recording and replay

The EEG (Fpz-Cz and Pz-Oz derivations), EOG (horizontal) and EMG (chin muscles) signals and a time code were FM-recorded (bandwidth DC-1250Hz, SNR≈44dB) on an instrumentation tape recorder. In order to synchronize the human and automatic classifications, start and stop times were also indicated

on the paper trace. Tape speed errors did not influence the results, because computer timing was synchronized to the time code during replay (see section VI.8.2). Paper speed errors were small (less than 15s accumulated after 2½h) and were corrected for by expanding or compressing the replayed data. All timing was performed with an accuracy of ≈1s.

VI.8.2. On-line preprocessing during tape replay

An overview of the EEG, EOG and EMG preprocessing during tape replay is presented in figure 7.

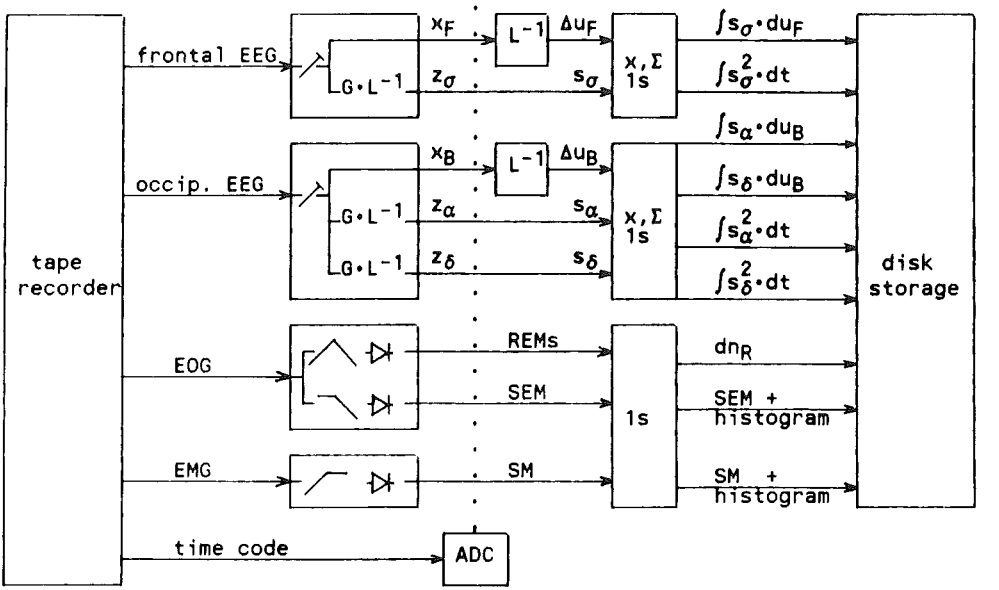


Figure 7: Block diagram of preprocessing the replayed analogue signals into digital data with a time resolution of 1s which are stored on disk. The symbols, $z_{.}$, correspond to those in figures 8 and 9, but not to those in chapter V.

Both EEG signals were fed into identical electronic circuits (fig.8), that carried out the following operations. Anti-anti-aliasing filtering (compensating the 70Hz lowpass filter that was applied in the

encephalograph), in order to reduce the bias mentioned in chapters V.4 and VI.1. Power calibration according to chapter VI.6. Reconstruction of the feedback signals, $s_\sigma(t)$, $s_\alpha(t)$ and $s_\delta(t)$, by $L^{-1}(f) \cdot G(f)$ according to figure 2 of chapter V.2.

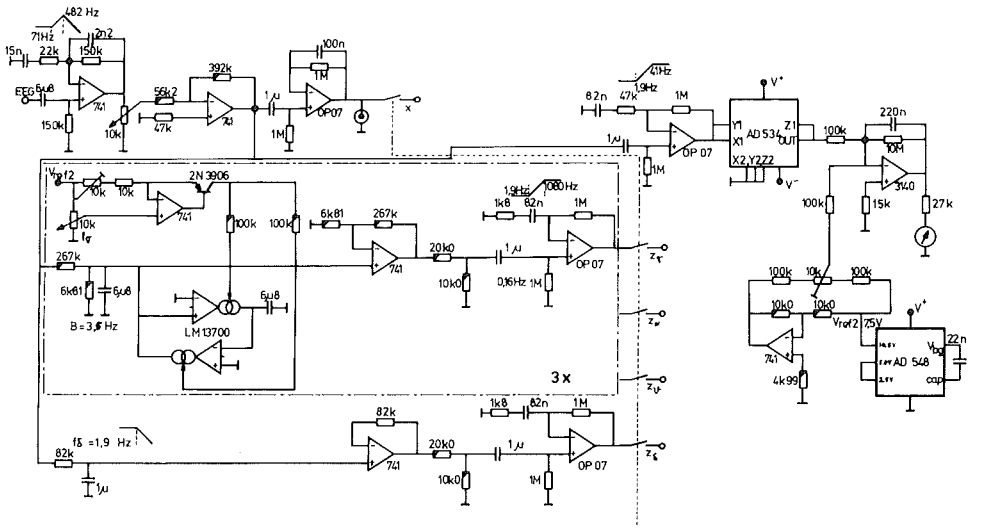


Figure 8: Analogue preprocessing of one EEG signal. Upper left; correction for 70Hz first-order lowpass filter in the encephalograph. The 10k potentiometer after this correction calibrates the signal power using the power measuring circuit on the right. Within dashed box; gyrator filter, G(f), with tunable centre frequency (10k potentiometer), fixed bandwidth (3.5Hz) and fixed gain (1). It is followed by the inverse-volume-conduction filter, $L^{-1}(f)$, resulting in the reconstructed feedback signals, $z_.$. The part within the dashed box is repeated 3 times in order to reconstruct sigma-, alpha- and theta-feedback signals. Below dashed box; reconstruction of delta-feedback signal, z_δ.

Another circuit (fig.9) measured the submental muscle tone, SM, using the EMG. It also detected rapid eye movements, REMs, and measured slow eye movements, SEMs, from the EOG.

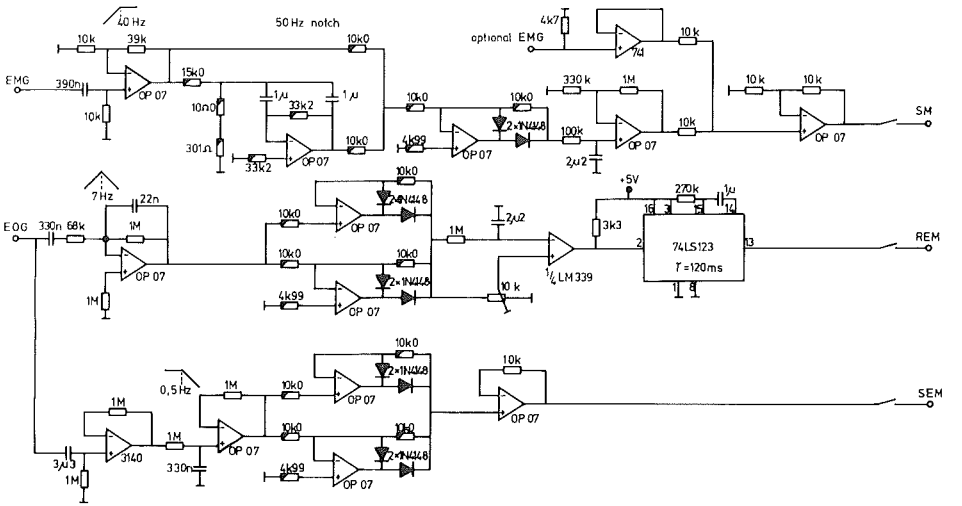


Figure 9: Analogue preprocessing of EMG and EOG signals. Upper part: highpass and notch filtering of the EMG for the purpose of EEG and hum rejection, followed by envelope measurement, resulting in the submental muscle tone, SM. Middle part: bandpass filtering, rectification and adaptive (noise-dependent) level detection of the EOG for the purpose of detecting REMs. Lower part: lowpass filtering and rectification of the EOG results in the slow eye movement amplitude, SEM.

The 8 resulting analogue signals were fed into a 12bit, 100Hz AD-converter, which was triggered and synchronized by the time code. This allowed accurate timing and synchronization to the human classification of the paper traces. The ADC performed part of the required scaling on the EEG signals (ch.VI.6). Therefore, only after this scaling, these signals correspond to the feedback signals, $s_{\cdot}(t)$. Digital reconstruction (according to chapter V.4) of the EEG signals by inverse-volume-conduction, L^{-1}, yielded the EEG observations, $\Delta u_F(t)$ and $\Delta u_B(t)$.

Subsequently, the cross products and squares that are required for the 'present' observations EEG likelihoods (ch.V.3 and V.4) were computed, using these reconstructed EEG observations and feedback signals. These products and squares were integrated over 1s. In order to detect REM-bursts, $du_R(t)=1$, rather than REMs, detected REMs were counted only if they were not preceded and followed by another REM within 3s. SEM amplitudes were averaged over 1s. If an SEM-maximum was detected, a whole-record SEM-peak histogram was updated. SM was averaged over 1s, resulting in the EMG-observations, $u_M(md)$. The whole-record SM histogram was updated every second.

The results had a time resolution of 1s and were stored on disk.

VI.8.3. Off-line processing

Adaptation to individual EOG and EMG levels

A high level, L_S, was chosen such that it was exceeded by only 9 SEM peaks from the whole-record SEM-peak histogram. Each upper crossing of the $L_S/2$ level was defined as a detected slow eye movement, $du_S(t)=1$.

The two mean muscle tone levels, μ_0 and μ_1, of figure 3 in chapter V.2.2 manifest as peaks that can be detected from the whole-record SM histogram. Such peaks did not occur (as was to be expected) in subjects who did not enter REM sleep. In these cases we chose a low level, L_M, such that the probability of it being exceeded by an EMG observation, $u_M(md)$, was $p=0.999$. We then set $\mu_0=1.06 \cdot L_M$ and $\mu_1=1.30 \cdot L_M$. These values correspond roughly to the locations of L_M, μ_0 and μ_1 in both subjects of the learning set (ch.V.2.2:fig.3).

Computation of likelihoods and optimal decision

Since all required parameters (μ_0, μ_1, a and the feedback gains, eye movement rates and muscle twitch probabilities indicated in table 1 of chapter V.2.2) and observations (EEG cross products and squares, REM-bursts, SEMs, SM) were now available, we computed the 'present' observations likeli-

hoods (ch.V.3 and V.4) with a time resolution of 1s. The EEG likelihoods were processed by the artifact rejection procedure described in chapter V.4.

Integration over D=30s yielded the 'present' observations likelihoods (see equation (17) in chapter V.3). These were stored in memory, which allowed execution of both the backward (ch.V.3:eq.(16)) and the forward (ch.V.3:eq.(15)) recursive algorithms that yielded the 'past' and 'future' observations likelihoods.

Multiplication (ch.V:eq.(14)) of 'past', 'present' and 'future' observations likelihoods then yielded the full observations likelihoods, $f_i(nD)$, of each sleep stage, i.

The optimal stage decisions on each time interval, $[nD,nD+D)$, are the ones that correspond to the largest likelihood, $f_i(nD)$.

VII. DISCUSSION

Previous work on automatic sleep stage classification involved many years of development and modifications. In spite of this, however, the results were not satisfactory (chapters I and V.1). The monitor, as described in this thesis, already seems to perform better at its first application (chapter V.5). This is probably due to some advantages of the model-based approach, that will be discussed now.

All previously reported systems start by preprocessing the EEG, EOG and EMG signals in such a way that partial decisions on isolated sleep-related signal characteristics are obtained. Such decisions may be 'sigma rhythm present/absent', 'REMs frequent/scarce', etc. These subdecisions are then combined, using a Boolean matrix or a decision tree, in order to obtain the final sleep stage classification.

The preprocessing methods result from a subjective choice between available signal analysis tools (like complex demodulation, Fourier analysis and Wiener filtering), that are sensitive to certain sleep-related signal characteristics (like sigma rhythm). The tools are generally imported from other applications and not specifically designed to cope with the particular aspects of our signals. This thesis shows that some preprocessors, that are based on specific models of the signal generating mechanisms, perform better, as will be discussed later.

Also, the combination of several subdecisions in order to obtain one final decision is suboptimal. Because of the probabilistic nature of the sleep-related signals, and their rather small 'signal-to-noise-ratio', false subdecisions (and even contradictory ones: e.g. the combination of subdecisions 'sigma rhythm present' and 'REMs frequent' does not correspond to any sleep stage) will often occur. False sleep stage classifications will

result. A better method would be able (as both the human procedure and the method designed in this thesis) to correct some of these by taking the reliability of the various subdecisions into account. For instance, very frequent REMs would dominate borderline sigma rhythm and prevail in the final stage classification. 'Fuzzy' systems methods (Kumar 1977) may account for this particular feature, but they are not tuned to the known probabilistic characteristics of sleep and sleep-related signals.

It has apparently been a problem to create methods that exploit this available knowledge about sleep more efficiently.

In this thesis, I developed new methods that are specifically designed to cope with well known characteristics of the sleeping system. The strategy for such a design was provided by model-based estimation and detection theory (chapters III, IV and V). This theory provides a framework that tells us precisely how to exploit the available knowledge of a system for the purpose of its analysis. The knowledge must be formulated, using mathematical, particularly probabilistic, models. A good model of the system results in the optimal analysis method. This approach had already proven its value in a simple application to the waking EEG (chapter VI.1). The combination of sleep and sleep-related EEG, EOG and EMG is more complex. However, I believe that the medical and physiological literature (chapters II, III, IV and V), supplemented with some own research (chapters II and V), provides enough material on the structure and statistics of sleep and sleep-related signals; enough material for the development of models that enable the design of better analysis tools (chapters II, III, IV, V and VI); and certainly too much material to ignore, as most of the previous work has done.

The theory handles stochastic models of sequential observations that

are mutually (conditionally) independent, for instance based on Markov chains, Wiener and Poisson processes. This corresponds to the highly unpredictable behaviour of sleep and sleep-related signals. However, the theory can only handle models with a relatively simple structure, as compared to the complexity of the sleeping system. Therefore, the available knowledge about sleep and sleep-related signals had to be simplified considerably in order to construct such models. While doing this, I had to make many subjective decisions, as mentioned in the corresponding chapters. The temporal lowpass filter as a model for the spatial lowpass volume conduction of the EEG (chapters III and V) and the Gaussian model for the EMG (chapter V) can still be improved considerably. Models for K-complexes and theta waves are currently being investigated.

Both the proposed Markov chain model for sleep (chapter II) and the white-noise-driven feedback model for the EEG (chapters III and V) lead to optimal analysis algorithms that differ largely from the commonly applied techniques. It was, therefore, reassuring to see that these models simulated sleep patterns and alpha rhythms, that could hardly be distinguished from real ones. This is not true for the EOG and EMG models. However, these do enable the monitor to exploit available statistical knowledge about EOG and EMG.

Once the system was modelled, the problem 'how to detect the sleep stages from the observed EEG, EOG and EMG signals' was reformulated in terms of the model. The simplifications proved to be sufficient: it was possible to derive the optimal solution to this problem in a form that enabled practical implementation. Some interesting and new signal analysis tools were developed. These are discussed here, and compared to previously applied methods.

The detection of the feedback gains of neural networks that generate

EEG rhythms (chapters III, V and VI.1) is probably more closely related to 'real' sleep, and performs better (ch.VI.1, Kemp 1983[b]) than the so far applied EEG preprocessors that are not based on physiological models. In particular, the previously applied EEG preprocessors are sensitive to the absolute power of EEG activities such as sigma rhythm and delta waves. This power is also largely influenced by age and by not sleep-related variables, such as electrode distance, morphology of the skull, etc. The results are, therefore, biased by these effects (e.g. Webb and Dreblow 1982), a drawback that is not shared by the feedback gain detector.

The exponential feedback integrators of chapters III, IV and VI.5, and the discrete time analogue in chapter V, provide an attractive alternative to the commonly applied segmentation approach. Segmentation (sometimes adaptive) was required in sleep classification algorithms, because sleep was considered to be a process with unpredictable non-stationarities. The problem was that on the one hand segments should be short in order not to miss small intervals, but on the other hand they should be long in order to minimize variance. Segmentation algorithms cannot account for stage transitions within segments, and supplementary measures must be taken in order to take the correlation between adjacent segments into account. Chapter II shows that the unpredictable sleep stage transitions can be approximately modelled as a stationary Markov process. The optimal compromise between long and short segments is therefore the corresponding (chapters III, IV, V and VI.5) stationary integrator with non-linear feedback; not segmentation. In this way, sleep can be classified continuously and each classification is based on weighed 'past', 'present' and 'future' information. Clear 'present' information (like a large sigma spindle that is reflected by a large gain detector output) increases the exponential feedback, thereby attenuating the influence of 'past' and 'future'. The signal-dependent temporal smoothing that is

applied by human classifiers performs comparably.

The forward and backward recursive vector algorithms in chapter V weigh (as humans) 'past', 'present' and 'future' of all sleep-dependent characteristics of all observed signals, before reaching the final decision on the present sleep stage. By this procedure, the monitor is capable of weighing contradictory information, something the decision trees used to date cannot do. A similar weighing algorithm has recently been applied to EEG only (Lacroix 1985), but it does not process 'future' information properly.

The subjective part of developing a sleep stage monitor has merely been shifted from selecting appropriate signal analysis tools towards selecting appropriate models. However, the feasibility of the approach is demonstrated by the first practical results. The man-machine agreement in a small group of both patients and healthy subjects was 70.6%: this is about 4.5% worse than the man-man agreement in this group (chapter V). This difference becomes 0.5% in favour of the automatic method if the, vaguely defined and diagnostically rather unimportant, difference between wakefulness and drowsiness (stage 1) is ignored (chapter V). The differences between the man-machine and man-man agreement so far reported in the literature for other systems (Johnson 1977, Lacroix and Stanus 1985[a,b]) are larger than those we found here. Therefore, this automatic sleep stage monitor seems to be an attractive alternative to the laborious, inconsistent, human procedure. It can and should be tested on more patient groups. Direct comparison to other automatic methods would be interesting.

Our hope was also to develop models which would be closer to real, physiological sleep. The detection of EEG feedback gains already accounts for the widely adopted assumption that EEG rhythms are generated by feedback

loops in the brain. Also, (unlike most other methods) the monitor does not require segmentation of the essentially continuous-time sleep process. However, both the human observer (acting according to the Rechtschaffen and Kales standard of 1968) and our models assume the existence of a limited number of discrete sleep stages. Although this seems to be a reasonable first approximation, some sleep related variables suggest a continuously varying sleep state or a division of stages into substates. This can be studied (chapter VI.2) using the unsmoothed, 'present' likelihoods of chapter V, since these are not based on this assumption but do exploit models of the sleep-related signals. From a physiological point of view, however, these models are rather superficial. Therefore, our methods for analysing sleep may still benefit substantially from specific research that links the observations to physiological functioning and to physiological mechanisms.

Using the model-based analysis methods and the 'Walkman' sleep recorder (chapter VI.4), we can now perform both sleep research and the diagnosis of sleep disturbances better, more consistently and more efficiently. This may result in improved understanding of sleep mechanisms and their defects. Furthermore, it may lead to the development of improved models. Because the developed algorithms are of sufficient generality to handle a variety of observations models, improved models may directly result in correspondingly improved analysers. Although it is still difficult to provide efficient therapy for most sleep complaints, it can now be evaluated more quickly, more consistently and in more cases.

REFERENCES

Alihanka J, Björkqvist SE, Oinaala M, Soini J, Valleala P. Oculomotor outflow and its relation to electrical activity of suprasylvian gyrus and pulvinar during paradoxical sleep in the cat. Electroencephalogr. Clin. Neurophysiol. 1975, 39:9-19

Anders TF, Keener M. Developmental course of nighttime sleep-wake patterns in full-term and premature infants during the first year of life I. Sleep 1985, 8:173-192.

Anders TF, Keener MA, Kraemer H. Sleep-wake state organization, neonatal assessment and development in premature infants during the first year of life II. Sleep 1985, 8:193-206.

Andersen P, Gillow M, Rudjord T. Rhythmic activity in a simulated neuronal network. J Physiol 1966, 185:418-428.

Aserinsky E. Rapid eye movement density and pattern in the sleep of normal young adults. Psychophysiol. 1971, 8:361-375

Bari F, Rubisek G, Benedek G, Obál F(jr), Obál F. Analysis of ultradian sleep rhythms in rats, using stage transition functions. Electroenc. Clin. Neurophys. 1981, 52:382-385.

Beersma DGM, Daan S, Van den Hoofdakker RH. Distribution of REM latencies and other sleep phenomena in depression as explained by a single ultradian rhythm disturbance. Sleep 1984, 7:126-136.

Boel R, Varaiya P, Wong E. Martingales on jump processes. II: Applications. SIAM J. Control 1975, 13:1022-1061

Borbély AA. A two process model of sleep regulation. Human Neurobiology 1982, 1:195-204.

Borbély AA. Sleep regulation: outline of a model and its implications for depression. In: Borbély AA, Valatx JL, eds. Sleep mechanics. Berlin: Springer Verlag, 1984: 272-284

Bowe TR, Anders TF. The use of the semi-Markov model in the study of the development of sleep-wake states in infants. Psychophysiology 1979, 16:41-48.

Březinová V. The number and duration of the episodes of the various EEG sta ges of sleep in young and older people. Electroenc. Clin. Neurophys. 1975, 39:273-278.

Broughton R. Polygraphic recordings of sleep and sleep disorders. In: Niedermeyer E, Lopes da Silva FH. eds. Electroencephalography. Baltimore, Munich: Urban and Schwarzenberg, 1982:571-598.

Broughton R, Green D, Sherif O, Mary J, Da Costa B. Automatic high speed detection of human sleep variables using a hardware pre-processor. Electroenc clin Neurophysiol 1979, 46:10P.

Broughton R, Healey T, Maru J, Green D, Pagurek B. A phase locked loop device for automatic detection of sleep spindles and stage 2. Electroenceph clin Neurophysiol 1978, 44:677-680.

Campbell K, Kumar A, Hofman W. Human and automatic validation of a phase-locked loop spindle detection system. Electroenceph clin Neurophysiol 1980, 48:602-605.

Campbell KB, Wilkinson RT. Sleep in the natural environment:physiological and psychological recording and analysing techniques. In: Biological rhythms, sleep and shift work (eds LC Johnson, WP Colquhoun, DI Tepas, MJ Colligan). In: Advances in sleep research 7 (ed ED Weitzman), 1981: 581-606.

Chouvet G, Odet P, Valatx J, Pujol J. An automatic sleep classifier for laboratory rodents. Waking and Sleeping 1980, 4:9-31.

Clements DJ, Anderson BDO. Well-behaved Itô equations with simulations that always misbehave. IEEE Trans AC 1973, 18:676-677.

Coates TJ, Killen JD, George J, Marchini E, Silverman S, Hamilton S, Thoresen

CE. Discriminating good sleepers from insomniacs using all-night polysomnograms conducted at home. J Nerv and Ment Dis 1982, 170:224-230.

Corner MA. Maturation of sleep mechanisms in the central nervous system. In: A Borbély, JL Valatx (eds). Sleep mechanisms. Berlin, Springer Verlag 1984.

Courtney P, Noton D. A hybrid computer system for unsupervised scoring of sleep records. Biomed Sci Instrum 1972, 9:161-168.

Daan S, Beersma DGM, Borbély AA. Timing of human sleep: recovery process gated by a circadian pacemaker. Am. J. Physiol. 1984, 246:R161-R178.

Davis MHA, Andreadakis E. Exact and approximate filtering in signal detec tion: an example. IEEE-IT 23, 1977: 768-772.

De Groen JHM, Koper H, Bergs PPE, Verheyen MJN, Caberg HB. Ambulatory sleep wake polygraphy in narcolepsy. Electroenceph clin Neurophysiol 1985, 60:420-422.

Dement W, Kleitman N. Cyclic variations in EEG during sleep and their rela tion to eye movements, body motility, and dreaming. Electroenc. Clin. Neurophys. 1957, 9:673-690.

Dumermuth G, Walz W, Scollo-Lavizarri G, Kleiner B. Spectral analysis of EEG activity in different sleep stages in normal adults. Eur Neurol 1972, 7:265-296.

Feinberg I, Koresko RL, Heller N. EEG sleep patterns as a function of normal and pathological aging in man. J. Psychiatric Res. 1967, 5:107-144.

Fleiss JL. Statistical methods for rates and proportions. New York: Wiley, 1981.

Gaillard JM, Blois R. Spindle density in sleep of normal subjects. Sleep 1981, 4:385-391

Gath I, Bar-On E. Computerized method for scoring of polygraphic sleep recordings. Computer Programs in Biomedicine 1980, 11:217-223.

Gath I, Bar-On E. Classical sleep stages and the spectral content of the EEG signal. Int J Neurosci, 1983, 22:147-156.

Goldberg RM, Beiber I. A portable REM-detecting machine. IEEE-BME 1979, 26:513-516.

Gopal IS, Haddad GG. Automatic detection of eye movements in REM sleep using the electrooculogram. Am J Physiol 1981, 241:R217-221.

Haralick RM. Decision making in context. IEEE-PAMI 1983, 5:417-428.

Hasan J. Differentiation of normal and disturbed sleep by automatic analysis. Acta Physiol Scand, supplementum 1983, 526:1-103.

Hasan J. Automatic analysis of sleep recordings: a critical review. Annals of Clinical Research 1985, 17:280-287.

Hauri P, Hawkins DR. Alpha-delta sleep. Electroenceph clin Neurophysiol 1973, 34:233-237.

Herman JH, Barker DR, Roffwarg HP. Similarity of eye movement characteristics in REM sleep and the awake state. Psychophysiol. 1983, 20:537-543

Hermann WM, Kubicki S. Various techniques of computer analysis in nocturnal sleep. In: Epilepsy, Sleep and Sleep Deprivation. eds R Degen, E Niedermeyer. Elsevier Science Publishers BV, 1984:207-229.

Hobson JA, McCarley RW, Wyzinski PW. Sleep cycle oscillation: reciprocal discharge by two brainstem neural groups. Science 1975, 189:55-58.

Hoffmann R, Moffitt A, Wells R, Sussman P, Pigeau R, Shearer J. Quantitative description of sleep stage electrophysiology using digital period analytic techniques. Sleep 1984, 7:356-364.

Inoue K, Kumamaru K, Sagara S, Matsuoka S. Pattern recognition approach to human sleep EEG analysis and determination of sleep stages. Memoires of the Faculty of Engineering, Kyushu University, 1982, 42:177-195.

Jacobs L, Feldman M, Bender MB. Eye movements during sleep. Arch. Neurol. 1971, 25:151-159

Jansen BH, Hasman A, Lenten R, Pikaar R. Automatic sleep staging by means of profiles. In: Medinfo 80, eds. Lindberg and Kaihara. North Holland Publish. Cy, 1980:385-389.

Jasper HH. The ten twenty electrode system of the International Federation. Electroenceph clin Neurophysiol 1958, 10:371-375.

Jayakar PB, Patrick JP, Sill J, Shwedyk E, Seshia SS. Artifacts in ambulatory cassette electroencephalograms. Electroenceph clin Neurophysiol 1985, 61:440-443.

Jazwinsky AH. Stochastic processes and filtering theory. New York, London. Academic Press, 1970.

Johnson LC. The EEG during sleep as viewed by a computer. In: EEG Informatics. A Didactic Review of Methods and Applications of EEG. ed A.Rémond Elsevier/North Holland Biomedical Press, 1977:385-406.

Kailath T. A general likelihood-ratio formula for random signals in Gaussian noise. IEEE-IT 1969, 15:350-361.

Kayed K, Hesla PE, Røsjø Ø. The actioculographic monitor of sleep. Sleep 1979, 2:253-260.

Kemp B. Optimal sigma sleep stage detectors. Electroenceph clin Neurophysiol 1981, 52:S15.

Kemp B. Optimal detection and filtering of a Markov brain state. In: Modelling and data analysis in biotechnology and medical engineering, pp 121-124. Vansteenkiste GC, Young PC, eds. Amsterdam, Ney York, Oxford: North-Holland 1983.

Kemp B. Accurate measurement of flash-evoked alpha attenuation. Electroen ceph clin Neurophysiol 1983, 56:248-253. (see also this thesis, ch.VI.1)

Kemp B. An optimal monitor for the rapid-eye-movement brain state. Biol Cybern 1986, 54:133-139. (also this thesis: chapter IV)

Kemp B, Blom HAP. Optimal detection of the alpha state in a model of the

human electroencephalogram. Electroenceph clin Neurophysiol 1981, 52: 222-225.

Kemp B, Jaspers P. Optimal detection of a finite-state Markov brain process, based on vector EEG observations. In: D Kleima (ed). Proc. Fifth Symp. on Information Theory in the Benelux, Enschede, 1984:102-108.

Kemp B, Jaspers P, Franzen JM, Janssen AJMW. An optimal monitor of the electroencephalographic sigma sleep state. Biol. Cybern. 1985, 51:263-270. (also this thesis: chapter III)

Kemp B, Kamphuisen HAC. Simulation of human hypnograms using a Markov chain model. Sleep 1986, 9:405-414. (also this thesis: chapter II)

Koffler DJ, Gotman J. Automatic detection of spike-and-wave bursts in ambulatory EEG recordings. Electroenceph clin Neurophysiol 1985, 61:165-180.

Ktonas PY, Smith JR. Automatic REM detection: modifications on an existing system and preliminary normative data. Int. J. Biomed. Computing 1978, 9:445-464

Kumar A. A real-time system for pattern recognition of human sleep stages by fuzzy system analysis. Pattern Recognition 1977, 9:43-46.

Lacroix B. Optimal Bayesian recursive filter applied to on-line scoring of polygraphic sleep recordings. In: Proc Mediterranean Electrotechnical Conference, Vol I, eds A. Luque, A.R. Figueiras Vidal, J.M.R. Delgado. Elsevier Science Publishers, 1985.

Lacroix B, Stanus E. Contribution à l'analyse automatique des stades du som meil. Leur identification en temp reel par traitement numérique de signaux électrophysiologiques. Hôpital Erasme Bruxelles, 1985.

Lacroix B, Stanus E. New algorithms for on-line automatic sleep scoring, and their application to mini and micro-computer. Journal A 1985, 26:91-97.

Lairy GC. Critical survey of sleep stages. In: Sleep 1967, 3d. Congr. Sleep Res. Basel: Karger, 1977:170-184.

Larsen LE, Walter DO. On automatic methods of sleep staging by EEG spectra. Electroenceph clin Neurophysiol 1970, 28:459-467.

Larson HJ, Shubert BO. Probabilistic models in engineering sciences II. New York, Chichester, Brisbane, Toronto: Wiley 1979.

Lawder RE. A proposed mathematical model for sleep patterning. J. Biomed. Eng. 1984, 6:63-69.

Lim AJ, Winters WD. A practical method for automatic real-time EEG sleep state analysis. IEEE-BME 1980, 27:212-220.

Liptser RS, Shiryayev AN. Statistics of random processes I: general theory. Berlin, Heidelberg, New York. Springer 1977.

Lopes da Silva FH, Van Rotterdam A, Barts P, Van Heusden E, Burr W. Models of neuronal populations: the basic mechanisms of rhythmicity. Prog Brain Res 1976, 45:281-308.

Martens WLJ, Declerck AC, Kums GJTM, Wauquier A. Considerations on a computerized analysis of long-term polygraphic recordings. In: H Stefan, W Burr (eds). Mobile long-term EEG monitoring. Gustav Fisher, Stuttgart/ New York, 1982:265-274.

Martin WB, Johnson LC, Viglione SS, Naitoh P, Joseph RD, Moses JD. Pattern recognition of EEG-EOG as a technique for all-night sleep stage scoring. Electroenceph clin Neurophysiol 1972, 32:417-427.

McCarley RW, Benoit O, Barrionuevo G. Lateral geniculate nucleus unitary discharge in sleep and waking: state- and rate-specific aspects. J. Neurophysiol. 1983, 50:798-818

McCarley RW, Hobson JA. Neuronal excitability modulation over the sleep cycle: a structural and mathematical model. Science 1975, 189:58-60.

McCarley RW. Mechanisms and models of behavioral state control. In: Hobson JA, Brazier MAB, eds. The reticular formation revisited. New York: Raven Press, 1980:375-403.

McPartland RJ, Kupfer DJ. Computerized measures of electrooculographic activity during sleep. Int. J. Biomed. Computing 1978, 9:409-419

Miles LE, Dement WC. Sleep and aging, chapters 1-8. Sleep 1980, 3:119-220.

Molinari L, Dumermuth G, Lange B. EEG-based multivariate statistical analysis of sleep stages. Neuropsychobiology, 1984, 11:140-148.

Monroe JL. Inter-rater reliability and the role of experience in scoring EEG sleep records: phase I. Psychophysiology 1969, 5:376-384.

Nelson JP, McCarley RW, Hobson JA. REM sleep burst neurons, PGO waves, and eye movement information. J. Neurophysiol. 1983, 50:784-797

Othmer E, Othmer SC. Electromyogram processing for sleep research. Int J Biomed Computing 1980, 11:33-39.

Parmeggiani PL, Morrison A, Drucker-Colin RR, McGinty D. Brain Mechanisms of Sleep: an overview of methodological issues. In: D.J. McGinty et al (eds). Brain Mechanisms of Sleep. Raven Press, New York, 1985:1-33.

Pivic RT, Bylsma FW, Nevins RJ. A new device for automatic sleep spindle ana lysis: the spindicator. Electroenceph clin Neurophysiol 1982, 54:711-713.

Rechtschaffen A, Kales A. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. Washington DC: Public Health Service, US Government Printing Office, 1968.

Riley TL, Peterson H. Sleep studies in the subject's home. Clin electroenceph 1983, 14:96-101.

Salzarulo P. Variations with time of the quantity of eye movements during fast sleep in man. Electroencephalogr. Clin. Neurophysiol. 1972, 32: 409-416

Scheuler W, Stinshoff D, Kubicki S. The alpha-sleep pattern. Neuropsychobiology 1983, 10:183-189.

Schiller F. Historical note on sleep and eye movements. Sleep 1984, 7:199-201

Sciarretta G, Bricolo A. Automatic detection of sleep spindles by analysis of

harmonic components. Med Biol Eng Comput 1970, 8:517-519.

Segall A. Stochastic processes in estimation theory. IEEE Trans. IT 1976, 22:275-286

Segall A, Davis MHA, Kailath T. Nonlinear filtering with counting observations. IEEE Trans. IT 1975, 21:143-149

Sewitch DE, Kupfer DJ. Polysomnographic telemetry using Telediagnostic and Oxford Medilog 9000 systems. Sleep 1985, 8:288-293.

Shouse MN, Sterman MB. Changes in seizure susceptibility, sleep time and sleep spindles following thalamic and cerebellar lesions. Electroenceph clin Neurophysiol 1979, 46:1-12.

Smith JR. Computers in sleep research. CRC Critical Reviews in Bioengineering 1978, 3:93-148

Smith JR. Automated analysis of sleep EEG data. In: Handbook of Electroencephalography and Clinical Neurophysiology. Revised Series, volume 2: Clinical Applications of Computer Analysis of EEG and other Neurophysiological Signals (eds. FH Lopes da Silva, W Storm van Leeuwen, A Rémond). Elsevier, Amsterdam, 1986.

Smith JR, Funke WF, Yeo WC, Ambuehl RA. Detection of human sleep EEG wave forms. Electroenceph clin Neurophysiol 1975, 38:435-437.

Snyder DL. Random point processes. New York: Wiley, 1975.

Spiegel R. Sleep and sleepnessness in advanced age. Lancaster: MTP Press 1981

Spreng LF, Johnson LC, Lubin A. Autonomic correlates of eye movement bursts during stage REM sleep. Psychophysiol. 1968, 4:311-323

Stefan H, Burr W (eds). Mobile long-term EEG monitoring. Gustav Fisher, Stuttgart / New York, 1982.

Sterman MB, Bowersox SS. Sensorimotor electroencephalogram rhythmic activity: a functional gate mechanism. Sleep 1981, 4:408-422.

Toth MF. A new method for detecting eye movement in sleep. Psychophysiology

1971, 7:516-523

Ursin R, Moses J, Naitoh P, Johnson LC. REM-NREM cycle in the cat may be sleep-dependent. Sleep 1983, 6:1-9.

Vaca MV, Snyder DL. Estimation and decision for observations derived from martingales: part II. IEEE Trans. IT 1978, 24:32-45

Van Schuppen JH. Filtering, prediction and smoothing for counting process observations, a martingale approach. SIAM J Appl Math 1977, 32:552-570

Van Trees HL. Detection, estimation and modulation theory.I. New York: Wiley 1979.

Vo-Ngoc B, Poussart D, Langlois JM. Automatic recognition of sleep spindles using short-term spectral analysis. Behav Res Meth Instrum 1971, 3: 217-219.

Webb WB, Dreblow LM. A modified method for scoring slow wave sleep of older subjects. Sleep 1982, 5:195-199

Williams RL, Agnew HW, Webb WB. Sleep patterns in young adults: an EEG study. Electroencephalogr. Clin. Neurophysiol. 1964, 17:376-381

Winfree AT. Human body clocks and the timing of sleep. Nature 1982, 297:23-27

Wong E. Recent progress in stochastic processes - a survey. IEEE Trans IT 1973, 19:262-275.

Wong E, Zakai M. On the relation between ordinary and stochastic differential equations. Int J Eng Sci 1965, 3:213-229.

Wonham WM. Some applications of stochastic differential equations to optimal non linear filtering. SIAM J Control 1965, 2:347-369.

Yang MCK, Hursch CJ. The use of a semi-Markov model for describing sleep patterns. Biometrics 1973, 29:667-676

Zung WWK, Naylor TH, Gianturco D, Wilson WP. A Markov chain model of sleep EEG patterns. Electroencephalogr. Clin. Neurophsyiol. 1965, 19:105

SUMMARY

Probabilistic models of sleep and of the sleep-related activity of brain, eyes and muscles have been proposed. These models describe how the corresponding overnight recordings of electroencephalogram (EEG), electrooculogram (EOG) and electromyogram (EMG) signals are influenced both by chance and by the various sleep stages. On the basis of these models, a signal analysis method has been developed that monitors how sleep travels through its various stages in the course of the night. In a group of 6 sleepers, the performance of this automatic method was in good accordance with the generally applied, laborious and subjective analysis by specialists.

The probabilistic generators in the models are Wiener, Poisson point and Gaussian processes and a Markov chain. The analysis algorithms, which are optimal for these models, have been developed using non-linear stochastic filtering theory and statistical detection theory. The analysis was realized using analogue electronic circuits and a computer programme.

Chapters II (the sleep process), III (EEG), IV (EOG) and V (EEG, EOG EMG, implementation and application) each contain their own specific summary. Chapter VI describes (a.o.) how most recordings can be performed unsupervised at the subjects homes, using a portable cassette recorder. Because this circumvents the need for hospitalization, ánd because the cassettes can be analysed automatically, the number of working hours required per 24h recording can be reduced from about 30 to 3.

EEN OP SLAAPMODELLEN GEBASEERDE SLAAPMONITOR

SAMENVATTING

Waarschijnlijkheidsmodellen van slaap en van de daarmee samenhangende aktiviteit van hersenen, ogen en spieren zijn opgesteld. Ze beschrijven hoe de corresponderende nachtelijke registraties van het electroëncefalogram (EEG), electroöculogram (EOG) en electromyogram (EMG) deels door toeval, maar ook deels door de diverse slaapstadia worden beïnvloed. Op basis hiervan werd een analysemethode ontwikkeld, die uit de geregistreerde aktiviteiten zo goed mogelijk opmaakt hoe de slaap in de loop van de nacht door de verschillende slaapstadia is gereisd. In een groep van zes slapers kwamen de uitkomsten van deze automatische methode goed overeen met die van de tot nu toe algemeen toegepaste, arbeidsintensieve en subjektieve analyse door specialisten.

De toevalsgeneratoren in de modellen zijn Wiener-, Poisson punt- en Gaussische processen en een Markov keten. De, voor deze modellen optimale, analysealgoritmen zijn ontworpen door middel van niet-lineaire stochastische filtertheorie en statistische detektietheorie. De analyse is gerealiseerd met analoge elektronika en een computerprogramma.

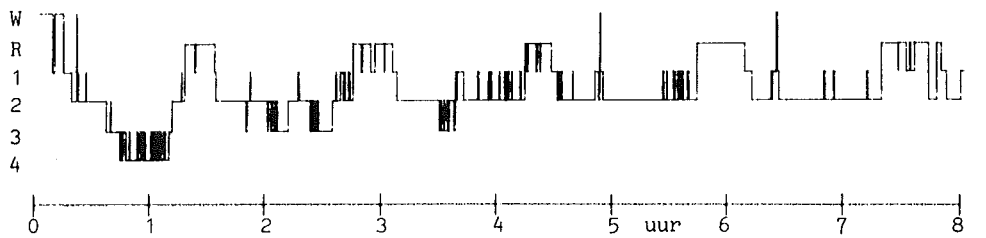
De hoofdstukken II (het slaapproces), III (EEG), IV (EOG) en V (EEG, EMG, EOG, implementatie en toepassing) bevatten elk hun eigen meer specifieke samenvatting. Hoofdstuk VI beschrijft o.a. hoe de meeste registraties bij de mensen thuis, zonder toezicht, kunnen worden uitgevoerd m.b.v. een draagbare cassetterecorder. Omdat ziekenhuisopname dan niet meer nodig is, én omdat de cassettes automatisch kunnen worden geanalyseerd, kan het aantal benodigde manuren per etmaalregistratie gereduceerd worden van ca. 30 tot ca. 3.

OVERZICHT VOOR NIET-VAKGENOTEN

De mens slaapt gedurende ongeveer 30% van zijn leven. Hoe we iemand ook proberen wakker te houden, iedereen slaapt binnen een paar etmalen in, vóór het tekort aan slaap zichtbare schade heeft aangericht. Het is daardoor niet duidelijk waarom we slaap zo hard nodig hebben.

Zo'n 10 à 20% van de mensen zegt niet goed te slapen en/of zich slaperig te voelen overdag. Dit kan tijdelijk zijn, maar soms gaat het om hinderlijke, chronische stoornissen, die zelfs levensbedreigend kunnen zijn. Bevredigende therapieën zijn nog schaars.

Tijdens slaap kunnen zes verschillende stadia worden onderscheiden. Zo herkennen we waak (W), doezelen (stadium 1), lichte slaap (stadium 2), diepe slaap (stadium 3), zeer diepe slaap (stadium 4) en droomslaap (ook wel REM-slaap (R) genoemd vanwege de snelle oogbewegingen (Rapid Eye Movements) die in dit stadium optreden). Goede slapers bewegen zich in de loop van de nacht volgens een soort golfbeweging door deze stadia, zoals te zien is in de slaapgrafiek van figuur 1. Deze grafiek wordt beïnvloed door lichamelijke en psychische faktoren, leeftijd, slaapstoornissen, slaapmiddelen, alcohol, nachtdienst, enzovoorts. Zo hebben ouderen bijvoorbeeld gemiddeld minder droomslaap dan kinderen.

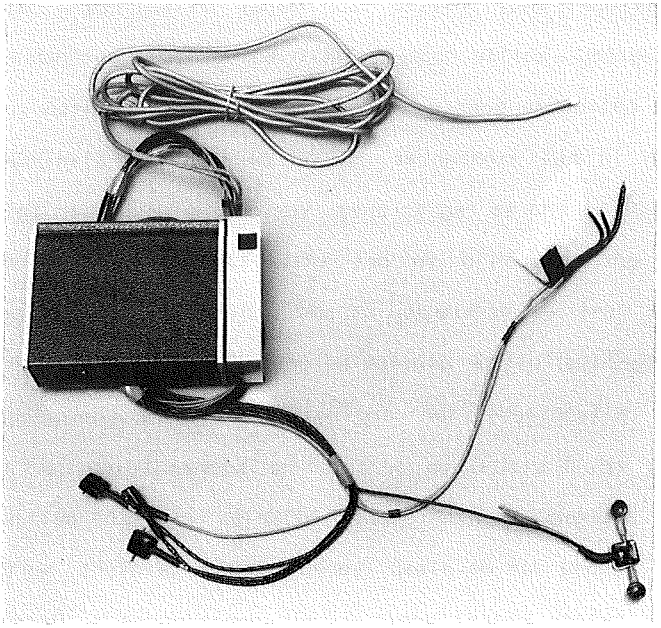


Figuur 1: Slaapgrafiek die laat zien hoe een gezonde slaper zich in de loop van de nacht door de slaapstadia W(aak), R(EM), 1, 2, 3 en 4 beweegt. In de eerste slaapgolf (van 0.12 tot 1.30 uur) daalt hij eerst via stadia 1, 2 en 3 af naar zeer diepe slaap (rond 1.00 uur), waarna de slaap weer minder diep wordt en om 1.15 uur de droomslaap (R) begint. Hierna volgen nog vier van zulke golven, die echter steeds minder diepe slaap bevatten.

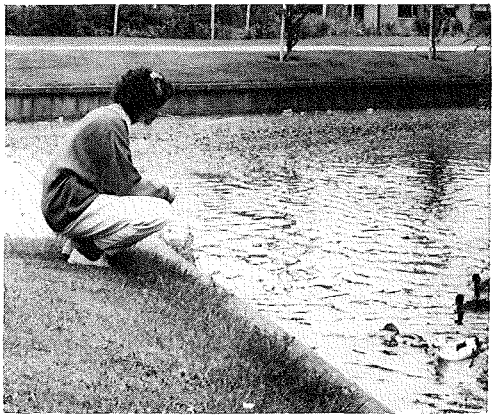
Zowel wetenschappelijk onderzoek naar het hoe en waarom van slaap, als de diagnostiek en therapie van slaapstoornissen, zijn meestal gebaseerd op slaapgrafieken. Om zo'n grafiek te kunnen maken, wordt gebruik gemaakt van het feit dat ieder slaapstadium zijn eigen, typische hersenaktiviteit (ElectroEncefaloGram: EEG), oogbewegingen (ElectroOculoGram: EOG) en spierspanning (ElectroMyoGram: EMG) heeft. Zo horen bijvoorbeeld bij droomslaap een onregelmatig EEG, snelle oogbewegingen en volledig ontspannen spieren. Proefpersonen en patienten moesten tot voor kort enige nachten (of etmalen) worden opgenomen in het ziekenhuis, om op de afdeling Klinische Neurofysiologie het EEG, EOG en EMG te registreren. Vervolgens werd door specialisten, die speciaal getraind zijn in het visueel herkennen van de verschillende slaapstadia, uit de verkregen strook papier (bijna 1km per etmaal) de slaapgrafiek afgeleid. Uit de grafiek werden vervolgens parameters zoals slaapstabiliteit, slaapdiepte en droomslaappercentage berekend. Registratie en verwerking van 1 etmaal op deze manier kosten beide ca. 15 manuur. Bovendien heeft ziekenhuisopname naast allerlei maatschappelijke problemen meestal ook tot gevolg dat de slaap, vooral de eerste nacht, wordt verstoord, zodat vaak een tweede etmaal moest worden geregistreerd. Daarom konden zowel wetenschappelijke research als patientenzorg slechts op zeer bescheiden schaal plaats vinden. Het effekt van slaapmiddelen werd nauwelijks objektief beoordeeld.

Het is nu mogelijk gebleken de meeste registraties zonder voortdurend toezicht bij de patient of proefpersoon thuis uit te voeren m.b.v. een 'Walkman'-achtige cassetterecorder (figuren 2, 3, 4, 5 en hoofdstuk VI.4). De onderzochte kan zich daarbij vrij bewegen in de eigen, vertrouwde omgeving. De kostbare ziekenhuisopname is dan niet nodig en de geregistreerde slaap is meer natuurgetrouw. Als de elektroden thuis worden aangebracht, kunnen tegelijkertijd een aantal omgevingsfaktoren worden waargenomen, die de slaapkwa-

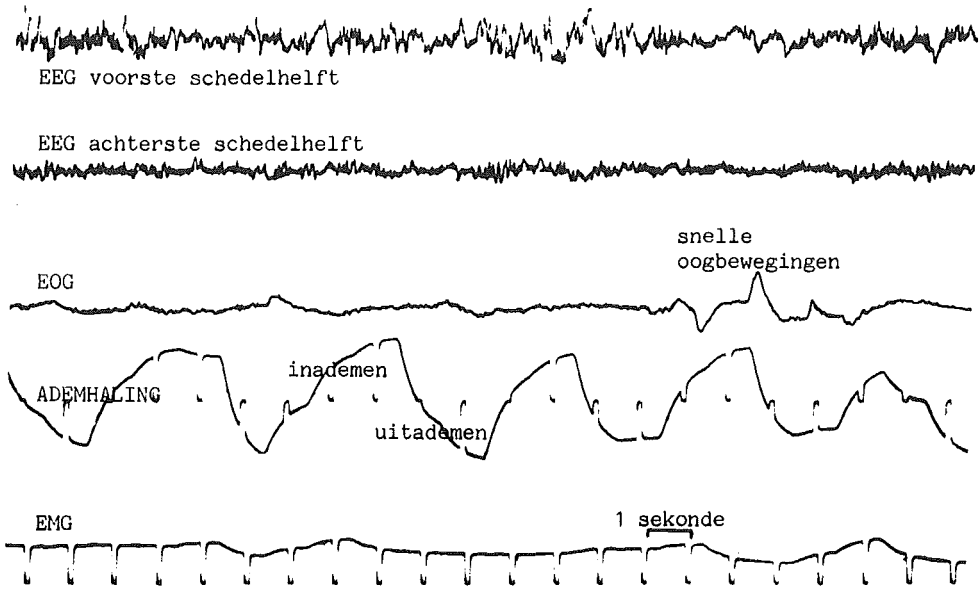
liteit mogelijk beïnvloeden. De commercieel verkrijgbare 4 kanaals recorder is door ons voorzien van extra elektronika, waardoor naast EEG, EOG en EMG ook nauwkeurige tijdinformatie, ademhaling (i.v.m. nachtelijke ademstilstand: bv. figuur 5) en lichaamstemperatuur (i.v.m. de 24-uurs bioritmiek) gedurende een heel etmaal op één standaard audio cassette kunnen worden vastgelegd.



Figuur 2: Draagbare slaaprecorder (links), met linksonder de EEG en EOG voorversterkers, rechtsonder de EMG voorversterker met elektrodes, rechtsboven de lichaamstemperatuur-sonde en midden-rechts de detector voor neus- en mondademhaling.

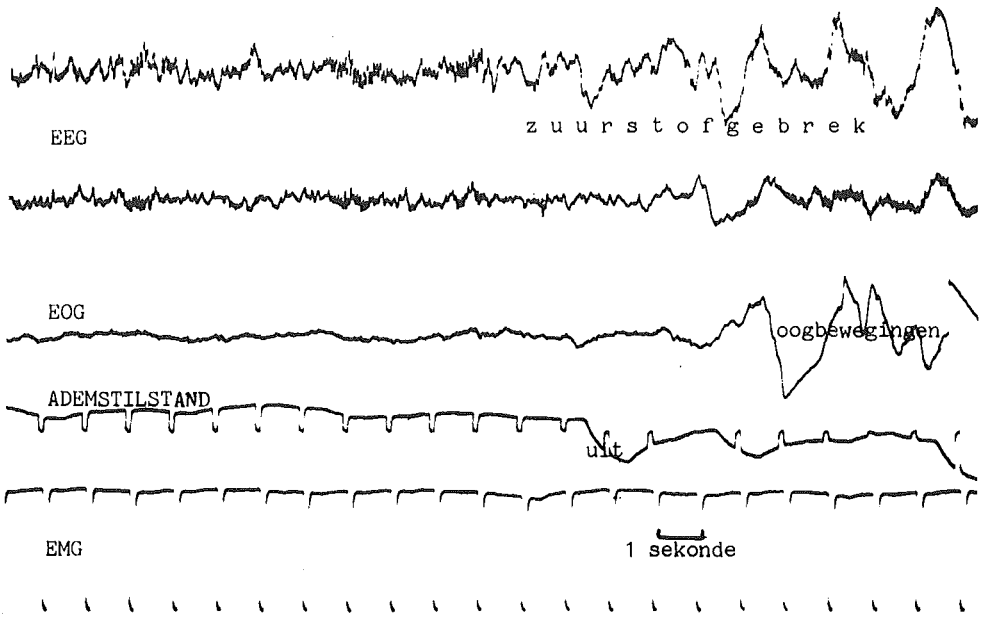


Figuur 3: Registratie van de elektrische aktiviteit van hersenen, ogen en kinspieren m.b.v. respektievelijk EEG-, EOG- en EMG-elektroden op de huid. De aktiviteiten worden vastgelegd op een draagbare (onder de kleding) cassetterecorder, zodat men zich vrij kan bewegen.

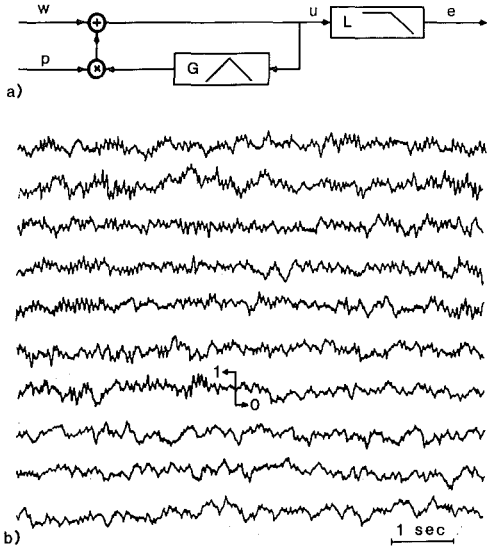


Figuur 4: Zo kunnen 22 seconden droomslaap er uit zien als de cassette op papier is afgespeeld. Let op de snelle oogbewegingen en het geringe EMG vergeleken met figuur 5.

Ook de tijdrovende verwerking van etmaalregistraties tot slaapgrafieken is geautomatiseerd. Hiertoe werd een computerprogramma ontwikkeld (hoofdstukken V en VI), dat de eigenschappen van EEG, EOG en EMG meet. Hoewel sommige eigenschappen kenmerkend zijn voor een bepaald slaapstadium, heeft ook het toeval een grote invloed. Figuren 4 en 5 illustreren dit: EEG en EOG veranderen van seconde tot seconde, ook al blijft het slaapstadium hetzelfde. Daarom gaat de computer uit van kansmodellen van hersenaktiviteit (hoofdstukken III,V), oogbewegingen (hoofdstukken IV,V) en spierspanning (hoofdstuk V). Deze modellen beschrijven hoe EEG, EOG en EMG deels door toeval, maar ook deels door de verschillende slaapstadia worden beïnvloed. Zo beschrijft bv. het EEG model (figuur 6) hoe de hersenen tijdens bepaalde slaapstadia meer of minder duidelijke elektrische trillingen produceren. De intensiteit van deze trillingen vertelt de computer dus iets over het slaapstadium.



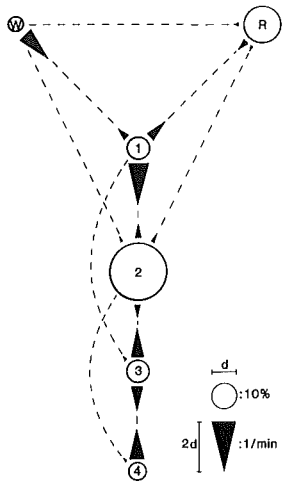
Figuur 5: Een ademstilstand (hier tijdens slaapstadium 2) veroorzaakt zuurstofgebrek en daardoor kort ontwaken. Bij deze patient gebeurde dit honderden keren in één nacht, zodat hij niet aan zijn diepe slaap toe kwam. Hij dacht zelf goed te slapen, maar klaagde over slaperigheid overdag.



Figuur 6a: Model van de elektrische aktiviteit van de hersenen. Het model wordt beïnvloed door slaap (p) en vormt ruisachtige aktiviteit (w) d.m.v. terugkoppeling (via G) om tot hersenaktiviteit (u), die via de schedel (L) op de hoofdhuid meetbaar is als EEG (e).

6b: Het met dit model nagebootste EEG bevat trillingen (als p=1, bv. tijdens stadium 2), die aan het eind verdwijnen (als p=0, bv. tijdens slaapstadium 1).

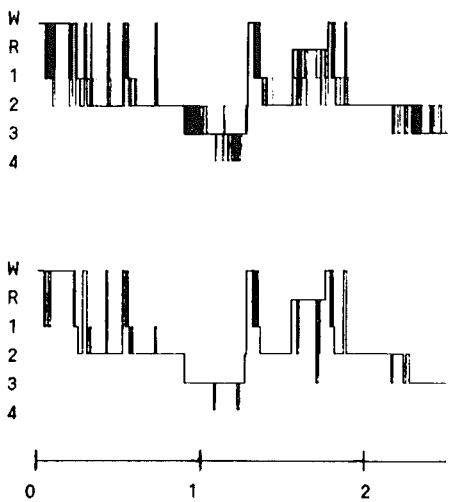
Op basis van de modellen 'weet' de computer welke eigenschappen toevallig zijn en welke kenmerkend zijn voor een bepaald slaapstadium. Hij kan dus op ieder moment op basis van het geregistreerde EEG, EOG en EMG inschatten wat het slaapstadium vermoedelijk is geweest. Omdat de invloed van het toeval groot is, kan het zo nu en dan lijken alsof de slaap naar een ander slaapstadium gesprongen is, terwijl dat in werkelijkheid niet zo is. Om dit zo goed mogelijk uit te kunnen maken, is ook een model van slaap zélf opgesteld (figuur 7 en hoofdstuk II). Dit model beschrijft hoe groot de kans op zo'n werkelijke sprong is. Bij kleinere sprongkansen zullen de EEG-, EOG- en EMG-eigenschappen duidelijker moeten zijn, voordat de computer beslist dat zo'n sprong waarschijnlijk toch is opgetreden.



Figuur 7: Model, dat beschrijft hoe slaap door de verschillende stadia (W, R, 1, 2, 3 en 4) reist. In ieder stadium is er konstant een kans dat naar een ander stadium wordt gesprongen. Grote sprongkansen corresponderen in de figuur met grote pijlen. Zo zal vanuit 1 meestal snel naar 2, maar soms ook naar W of R worden gesprongen.

De computer begint, net als een getrainde specialist, met het meten van slaapkenmerken zoals de genoemde EEG trillingen, snelle oogbewegingen, enz. Vervolgens worden deze kenmerken tegen elkaar afgewogen (hoe duidelijker het

kenmerk, hoe zwaarder het telt) om iedere ½ minuut tot een beslissing te komen in welk stadium de slaap zich vermoedelijk bevond. Door de grote invloed van het toeval zijn de echte slaapkenmerken soms nauwelijks te onderscheiden. Daarom is de beslissing moeilijk en zal de computer het niet altijd eens zijn met de specialist. Om dezelfde reden zijn ook twee specialisten het niet altijd met elkaar eens. De computermethode heeft het voordeel van consistentie: herhaalde analyse van dezelfde registratie geeft steeds dezelfde slaapgrafiek, terwijl zelfs een intensief getrainde specialist steeds tot verschillende resultaten komt. Uit toepassing op 6 slapers blijkt dat computer en specialist het gedurende 29% van de nacht oneens zijn over het juiste stadium, terwijl dat voor twee specialisten onderling 25% is (figuur 8). Het verschil tussen beide percentages is kleiner dan de 8% die tot nu toe van andere systemen bekend is. Dat komt waarschijnlijk doordat in dit proefschrift voor het eerst is gekozen voor een volledig op slaapmodellen gebaseerde aanpak.



Figuur 8: Slaapgrafieken van één korte (2½ uur) registratie, zoals verkregen door zes getrainde specialisten (boven; over elkaar heen getekend) en door de computer (onder).

Als het, uit diagnostisch oogpunt niet erg belangrijke, verschil tussen waak en doezel buiten beschouwing wordt gelaten, wordt de overeenkomst tussen de computer en een specialist zelfs iets groter dan die tussen twee specialisten onderling. De computeranalyse lijkt dus een goed alternatief te bieden voor de subjektieve methode. Er moet nog wel meer ervaring mee worden opgedaan.

Met de kombinatie van thuisregistratie m.b.v. de cassetterecorder en computeranalyse van de cassette kan het aantal benodigde manuren per etmaalregistratie gereduceerd worden van ca. 30 tot ca. 3 (exclusief eventuele reistijd bij het thuis aanbrengen van de elektroden). Momenteel verkrijgbare kleine computers kunnen een etmaalcassette in ongeveer 2 uur verwerken. Door concentratie in een klein aantal slaapcentra hoeven de apparatuur- en materiaalkosten daarbij niet toe te nemen. Slaapdiagnostiek en therapie kunnen dan zowel direkt als via sneller toenemende wetenschappelijke kennis verbeteren.

Stellingen bij het proefschrift: MODEL-BASED MONITORING OF HUMAN SLEEP STAGES

1. Door middel van slaapregistratie aan huis en computeranalyse van deze registraties kunnen de kosten van diagnostisch slaaponderzoek met een faktor tien worden gereduceerd.

   Dit proefschrift.

2. Het huidige, internationaal toegepaste standaardmodel van slaap kan verbeterd worden op grond van de huidige kennis.

   Dit proefschrift.

3. De in de techniek ingeburgerde toepassing van een anti-terugvouwfilter (volgens Nyquist) vóór signaalbemonstering staat op gespannen voet met het mathematisch aantrekkelijke gebruik van stochastische Itô-integraal vergelijkingen.

   Dit proefschrift.

4. Sommige recent met behulp van stochastische filtertheorie afgeleide analysemethoden kunnen ook worden verkregen via de klassieke maximum-likelihood aanpak.

   Dit proefschrift.

5. De ondoorgrondelijkheid en onvoorspelbaarheid van de werking van de hersenen nopen tot het gebruik van struktureel eenvoudige waarschijnlijkheidsmodellen.

6. Het beloop van ziekten, die gepaard gaan met complexe ongewilde bewegingen, kan eenvoudig worden gekwantificeerd met Doppler radar technieken.

7. Aan het formuleren van een probleem wordt vaak ten onrechte minder aandacht besteed dan aan het oplossen ervan.

8. Gedeeltelijke overlap van medische en technische opleidingen maakt zowel de verschillen van inzicht als het door de opgeleiden gezamenlijk bestreken vakgebied kleiner, waardoor de creativiteit in het samenspel tussen beide disciplines niet wordt bevorderd.

9. De afweging van kosten en baten van milieuvervuilende industrie kan worden bevorderd door de schoorsteenhoogtes te beperken, zodat de vervuiling bij de bron wordt geconcentreerd. De naleving van voorschriften omtrent deze hoogtes is eenvoudig controleerbaar.

10. Promotiebanen dreigen promoveren soms tot doel van onderzoek te promoveren.

11. 'Artificially intelligent expert systems' dienen onder toezicht te staan van intelligente experts.

B. Kemp **Hazerswoude, 11 juni 1987**