

# EEG analysis with nonlinear deterministic and stochastic methods: a combined strategy

Jürgen Fell<sup>1</sup>, Alexander Kaplan<sup>2</sup>, Boris Darkhovsky<sup>3</sup> and Joachim Röschke<sup>1</sup>

<sup>1</sup>Department of Psychiatry, University of Mainz, Untere Zahlbacherstr. 8, D-55101 Mainz, Germany; <sup>2</sup>Department of Human Physiology, Moscow State University, 119899 Moscow, Russia; <sup>3</sup>Institute for Systems Analysis, Russian Academy of Sciences, Prospect 60-letya Oktyabria, 117312 Moscow, Russia

Review

**Abstract.** We describe nonlinear deterministic *versus* stochastic methodology, their applications to EEG research and the neurophysiological background underlying both approaches. Nonlinear methods are based on the concept of attractors in phase space. This concept on the one hand incorporates the idea of an autonomous (stationary) system, on the other hand implicates the investigation of a long time evolution. It is an unresolved problem in nonlinear EEG research that nonlinear methods *per se* give no feedback about the stationarity aspect. Hence, we introduce a combined strategy utilizing both stochastic and nonlinear deterministic methods. We propose, in a first step to segment the EEG time series into piecewise quasi-stationary epochs by means of nonparametric change point analysis. Subsequently, nonlinear measures can be estimated with higher confidence for the segmented epochs fulfilling the stationarity condition.

**Key words:** EEG, nonlinear dynamics, stationarity, segmentation, change point analysis

## NEUROPHYSIOLOGICAL BACKGROUND

### Single cell level

Should the activity of a single neuron be treated in a deterministic or in a stochastic manner? A stochastic formalism is often used to model the behavior of a system consisting of a large number of independent, conceptually deterministic elements (e.g. in thermodynamics). In this case a deterministic description would fail because of the large number of variables and the incomplete knowledge of all initial conditions. The technical issue related to the appropriateness of stochastic *versus* deterministic modelling has to be distinguished from ontological reflections regarding the dynamical nature of a phenomenon. From the ontological perspective, it is clear that "true" randomness only occurs in the domain of quantum mechanics, where an atom or molecule can be described by a superposition of different basis states, which are observable with different probabilities. In this sense, all physical systems are probabilistic in nature. But in order to have a macroscopic impact, a quantum mechanical system in a superposed state has to be separated from macroscopic environment and the outcome of the observation (measurement) has to be amplified into the macroscopic domain (like in Schrödinger's famous Gedankenexperiment). On the other hand, large numbers of ion channels respectively ions are engaged in the generation of action potentials. It was for example estimated, that nonmyelinated axons depending on the cell type possess 35-500  $\text{Na}^+$  channels per  $\mu\text{m}^2$  with ionic currents in the range of  $10^7$  ions/s (Koester 1991). Synaptic transmission relies on the exocytosis of vesicles filled with several thousands of neurotransmitter molecules. Therefore, randomness originating from isolated quantum mechanical objects is not to be expected to have an influence on neural dynamics (unless a coherent quantum field would be discovered coupling all ions respectively neurotransmitter molecules together, which is unlikely).

But also from a technical point of view a deterministic approach has proven to be successful to model neuroelectric activity on the single cell level. Neural membrane potentials result from intra- and extracellular ion concentrations (e.g.  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$ ). The ion concentrations are mainly ruled by complex interactions of voltage-gated ion channels. This means, neuroelectric dynamics essentially are generated by a mechanism of

nonlinear reinforcement (membrane potential  $\leftrightarrow$  voltage-gated ion channels). Already in 1952 Hodgkin and Huxley suggested a nonlinear deterministic model for the electric activity of squid axons. The model consisted of four coupled first-order differential equations, where sodium and potassium activation, sodium inactivation and the membrane potential (respectively the membrane current in case of voltage clamp) were implemented as variables. Dynamical features similar to those predicted by the Hodgkin-Huxley model were experimentally observed in periodically stimulated squid axons (e.g. Aihara and Matsumoto 1986, Matsumoto et al. 1987) and in mollusc neurons (e.g. Hayashi et al. 1982, Holden et al. 1982). A footprint of nonlinear deterministic dynamics are characteristic transitions (bifurcations) from periodic to chaotic dynamics such as period doubling and intermittency. Bifurcations were indeed detected for periodically driven squid and molluscan neurons (e.g. Hayashi et al. 1983, Aihara and Matsumoto 1987). These findings point to the view that neuroelectric activity on the single cell level can successfully be treated by a nonlinear deterministic approach.

### EEG dynamics

Scalp recorded EEG signals mainly result from excitatory postsynaptic potentials of cortical pyramidal cells (see e.g. Creutzfeld 1983). In order to give rise to a recognizable EEG component, the synchronized activity of some ten thousands of neurons is necessary. Cortical pyramidal cells are driven by thalamic nuclei, which themselves are under influence of mesencephalo-reticular input. Besides the external dynamical pacing of cortical delta, theta and alpha rhythms, there is evidence for autosynchronized cortical rhythms at higher frequency ranges. One example is the sensory evoked gamma activity, which was suggested to play a major role in primary sensory processing and feature binding (e.g. Gray et al. 1989). In contrast to intracortically recorded EEG signals, the topography of scalp recorded EEG components is smeared out by the resistive properties of the skull. Since this is not the case for MEG data, magnetoencephalography today is preferred for tasks, which require a high localisation accuracy (e.g. Hari et al. 1989).

In the previous paragraph we argued that single neurons are highly nonlinear elements. On the neuronal group level, more nonlinearity is introduced through multiple feedback loops at each of the hierarchic levels

of cortical processing. An impressive overview of the distributed and mutually interconnected processing in the visual system of primates for example can be found in the article of Felleman and Van Essen (1991). Clear evidence for nonlinearity is indeed detectable in the scalp registered EEG. Nonlinear systems, which are driven by a periodic signal do not only show a response at the driving frequency (like linear systems would do), but also at harmonics and subharmonics (e.g. Lauterborn and Holzfuss 1991). This effect can be recognized under photic or auditory driving conditions, which means periodic visual or auditory stimulation with a frequency in the range of dominant EEG rhythms. Hereby, the so-called steady-state EEG response consists of the stimulating frequency, but also of harmonic and subharmonic components (e.g. Jansen 1986).

Schiff and coworkers (1994) demonstrated, that the electric activity of an assembly of hippocampal pyramidal neurons can be converted into periodic rhythms by applying a method developed for control of chaotic systems (Ott et al. 1990). This controlling chaos technique relies on characteristic features of deterministic chaotic systems, such as that dynamical states in phase space converge to an attractor and that there are so called stable and unstable directions, where states are attracted to respectively repelled from the attractor. On the other hand, it was shown that the same chaos control technique works also in some simple stochastic, but non-chaotic systems (e.g. Christini and Collins 1995), so that evidence for nonlinear deterministic dynamics in hippocampal neurone assemblies is not conclusive. Scalp recorded EEG can be regarded as a summation of topographically transformed intracortical EEG components. Therefore, based on the arguments for the single cell level a nonlinear deterministic model would anyhow be well founded. Nevertheless, from a more technical point of view the question remains whether activity generated by such a large number of quasiindependent elements may best be treated by a nonlinear deterministic or a linear respectively nonlinear stochastic approach.

In several recent investigations, dynamical features of EEG signals in terms of nonlinear measures were compared with phase randomized control data, which are in accordance with a linear stochastic model, but possess the same power spectra as the original data. Hereby, small but systematic deviances of the original EEG characteristics from the surrogate data were found leading to a statistically significant differentiation (e.g. Pritchard et al. 1995, Rombouts et al. 1995, Palus 1996). For sleep

EEG data, lower correlation dimensions and Lyapunov-exponents were estimated for the original EEG signals compared to phase randomized data, although the time courses during the night appeared to be similar (Achermann et al. 1994, Fell et al. 1996). With respect to saturation of the correlation dimension, all authors concluded that no low-dimensional dynamics are present. From these results, it follows that EEG signals may in first order be treated by a linear stochastic approach, but that a complete description might only be given by a high-dimensional nonlinear deterministic or a nonlinear stochastic model. From this perspective both, stochastic methods, which are based on the theory of linear stochastic systems, and nonlinear methods, which are suitable for the analysis of low-dimensional deterministic systems, appear equally appropriate (respectively non-appropriate) for the analysis of EEG signals.

## THE NONLINEAR DETERMINISTIC APPROACH: CONCEPTS, METHODS AND FINDINGS

### Concepts

The theory of nonlinear dynamical systems deals with deterministic systems, meaning systems where - in contrast to stochastic processes - the future development is totally determined by the momentary state. Typically, nonlinear dynamical systems are described by a set of  $n$  differential equations,  $n$  being the number of variables respectively components of the so-called state vector. A  $n$ -th order continuous-time dynamical system can be defined by the state equation:

$$\dot{x}(t) = f(x, \mu, v(t), t), \quad x(t) \in \mathbb{R}^n$$

where  $(t)$  is the time derivative of  $x(t)$ .  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  is called vector field and  $\mu$  and  $v(t)$  are the so-called control parameters of the system. That means, changes of the state vector  $x(t)$  are determined by a time-dependent function of  $x(t)$  itself and the control parameters. If the system is not explicitly time dependent, that is if

$$\dot{x} = f(x, \mu),$$

the system is called autonomous. Thus, in case of autonomous dynamics changes of  $x(t)$  are determined by a time-independent function of  $x(t)$  and also the control parameters are independent of time. If the state equation

is nonlinear, the superposition principle of linear systems does not hold true, that is

$$f(\lambda(x + y)) \neq \lambda f(x) + \lambda f(y),$$

The state equation of a nonlinear deterministic system as a rule can not be solved analytically by a finite number of integrations. Generally, only approximative numerical integration methods (e.g. Runge-Kutta-algorithms) can supply information about solutions of such systems.

An important tool for the investigation of dynamical systems (especially of the nonlinear autonomous kind) is the description in phase space. The phase space is defined by the set of all states, which can be reached by a certain class of systems. Phase spaces are usually differentiable manifolds, thus they can be treated locally like Euclidian spaces  $\mathbb{R}^n$ . The flow  $\Phi$  corresponding to a vector field  $f$  is a transformation, that maps the initial state  $x_0$  to a subsequent state  $x_t$ . The set of states  $\Phi_t(x_0)$ , reached by the system coming from a certain initial state  $x_0$  is called trajectory. So-called limit sets give information about the long time behavior of trajectories in phase space. Typical limit sets of conservative systems (energy preserving systems) are limit cycles in case of periodic motion and  $k$ -tori for quasiperiodic motion consisting of a superposition of several indecomposable frequencies. For dissipative systems (non energy preserving) the volume occupied by the states in phase space shrinks as time goes towards infinity (Theorem of Liouville). The limit set of an autonomous dissipative system to which trajectories converge for time increasing towards infinity is called attractor. It is obvious that - beside limit cycles and  $k$ -tori - equilibrium points are attractors corresponding to dissipative systems. But there is another less obvious possibility. In case of an at least third order nonlinear dissipative system the trajectories may be attracted by a limit set with a complicated geometrical structure. Because of their strange scaling behaviour (scaling invariance) similar to the structure of fractal sets (e.g. the Cantor-set) these limit sets were named strange attractors. They are footprints of deterministic chaos. Nevertheless, also strange nonchaotic attractors exist (e.g. Heagy and Hammel 1994) so that evidence for deterministic chaos in the presence of a strange attractor is not conclusive.

The reason for the strange phase space structure of chaotic attractors is the so-called sensitive dependence on initial conditions of chaotic systems. This means that two trajectories being infinitesimally close initially will

become exponentially separated at a rate characteristic for the system. In other words, only identical initial states result in the same final states ("weak causality"). Therefore, chaotic systems can not be predicted over longer periods of time. The unpredictability of chaotic motion is a gradual phenomenon meaning the reliability of predictions (in terms of probabilities) decreases exponentially with time. Hence, a certain time limit exists for every chaotic system leading to unpredictability for practical reasons (rounding errors of computers, observation errors etc.).

Small alterations of a control parameter typically result in small quantitative changes of a nonlinear system's dynamic. However, at some values, where the system is structurally unstable, a small variation of the control parameter will provoke a qualitative change of the dynamical behavior. This phenomenon is called bifurcation. Up to now, three different bifurcation routes to chaos are known. The period-doubling route can be observed for example in the well-studied logistic map. In this case, limit cycles switch to limit cycles with a period twice as long until at a critical value of the control parameter chaos occurs (Feigenbaum 1978). The intermittency route is characterized by stochastic alteration of periodic and irregular motion, whereby irregular time intervals increase, when the control parameter is turned towards a critical value (Manneville and Pomeau 1979). Finally Newhouse, Ruelle and Takens (1978) have shown, that starting from a fixed point a transition from quasiperiodicity to chaos is already possible after three "Hopf-bifurcations", each of which introduces a new fundamental frequency into the system (Ruelle-Takens-Newhouse-route).

A common problem concerning the investigation of high dimensional systems in reality is, that only one observable is known by measurement. The theorem of Takens (1981) deals with this situation. Takens has proven, that, if  $f$  is a smooth ( $\in C^{\geq 2}$ ) vectorfield on the  $m$  dimensional manifold  $M$  generating the flow  $\Phi$  and  $y$  a smooth measurement function,

$$E_{f,y} : M \rightarrow \mathbb{R}^{2m+1}$$

$$E_{f,y}(x) = (y(\Phi_0(x)), y(\Phi_\tau(x)), y(\Phi_{2\tau}(x)), \dots, y(\Phi_{2m\tau}(x)))$$

is generally an embedding of the original attractor. Embedding essentially means, that original and reconstructed attractor have the same topological properties. The  $2m+1$  components of the state vectors  $\tilde{x}$  in embed-

ding space simply result from the measured values  $y_t$  time shifted about the "delay time"  $\tau$ , that is  $\tilde{x}_t = (y_t, y_{t+\tau}, y_{t+2\tau}, \dots, y_{t+2m\tau})$ . According to Whitney's theorem the minimal dimension of the embedding space is  $2m$ ,  $m$  being the lowest possible dimension of a manifold that contains the original attractor. Takens (1981) could show that embedding into  $\mathbb{R}^{2m+1}$  is generally sufficient.

Although the time-delay method is independent of the choice of  $\tau$  on principle, the reconstruction may become unfavorable for practical reasons. A too small  $\tau$  yields  $\Phi_{k\tau}(x) \cong \Phi_{(k+1)\tau}(x)$  and the reconstructed attractor is "compressed" to the diagonal of the embedding space. On the other hand, for a too large time delay  $\Phi_{k\tau}(x)$  and  $\Phi_{(k+1)\tau}(x)$  become dynamically unrelated and the structure of the attractor disappears. Moreover, if  $\tau$  is close to a periodicity of the system, the periodical component will be under-represented in the reconstruction (Parker and Chua 1989). As long as these extreme situations are avoided, the exact value of the delay time  $\tau$  is almost arbitrary. Common choices are the first zero crossing of the autocorrelation function or the first minimum of the mutual information. These selections insure that two successive delay coordinates in some sense are as independent as possible.

A useful concept for the classification of fractal objects like chaotic attractors is given by the generalized Renyi-dimensions (Renyi 1970), which in contrast to the topological dimension of manifolds allow noninteger values. These measures yield a lower boundary of the degrees of freedom meaning the number of independent variables of a system. The Renyi-dimension is often called complexity of a system, although this is not a uniquely specified term and other definitions exist, for example algorithmic complexity (Chaitin 1975). Periodic and quasiperiodic systems are characterized by integer dimensions, whereas the phase space structure of chaotic systems implicates fractal (noninteger) dimensions. The Renyi-dimension easiest to calculate numerically is the correlation dimension  $D_2$ . Let the attractor be covered by  $N$  volume elements with diameter  $\epsilon$ . The probability for a system state to be located in the volume element  $i$  be  $P_i$ . Then  $D_2$  is defined by:

$$D_2 = \lim_{\epsilon \rightarrow 0} \left( \log \sum_{i=1}^{N(\epsilon)} P_i(\epsilon)^2 / \log \epsilon \right),$$

This expression is equivalent to

$$D_2 = \lim_{\epsilon \rightarrow 0} (\log C(\epsilon) / \log \epsilon)$$

where:

$$C(\epsilon) = \lim_{N \rightarrow \infty} \left( \sum_{j=1; i < j}^{N(\epsilon)} \Theta(\epsilon - \|x_i - x_j\|) \right),$$

$$\Theta(y) = 1 \text{ for } y > 0, \Theta(y) = 0 \text{ otherwise.}$$

$x_i$  respectively  $x_j$  are the state vectors in the embedding space. The quantity  $C(\epsilon)$  is called correlation sum.  $C(\epsilon)$  is just the number of pairs of points (state vectors), whose distance in phase space is closer than  $\epsilon$ . The scaling behavior in phase space is characterized through the relationship  $C(\epsilon)$  being proportional to  $\epsilon^{D_2}$ . Thus, the correlation dimension  $D_2$  finally results from the slope of the curves  $\log C(\epsilon) / \log \epsilon$ .

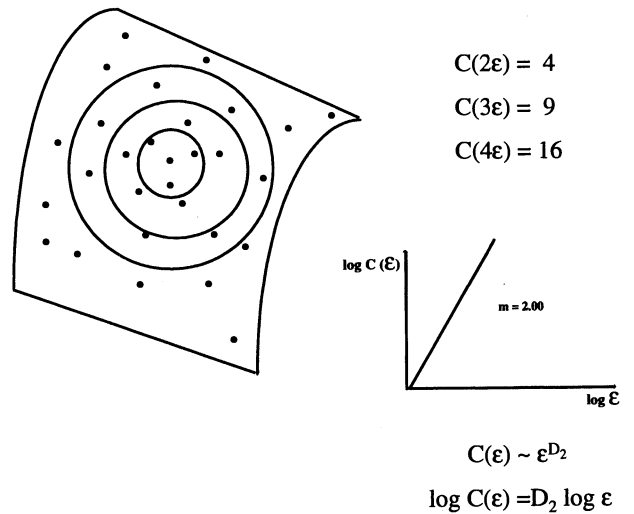


Fig. 1. Exemplarily a two dimensional manifold is shown. The algorithm is counting the number of points inside a circle with radius  $\epsilon$ . Plotting  $\log C(\epsilon)$  vs.  $\log \epsilon$  a straight line with slope  $m = 2$  occurs.

The sensitive dependence on initial conditions of chaotic systems can be quantified by the so-called Lyapunov-exponents. These values measure the mean exponential divergence or convergence of nearby trajectories in phase space. The evolution of an initial point  $x = x_R + \epsilon$  close to a point  $x_R$  on a reference trajectory can be approximated by the linearized flow map  $D_{x_R} \Phi_t(x_R)$ :

$$x(t) \approx x_R(t) + D_{x_R} \Phi_t(x_R) \epsilon$$

Lyapunov-exponents are defined by the logarithms of the eigenvalues  $\lambda_i(t)$  of the linearized flow map divided through time for times increasing towards infinity:

$$Li = \lim_{t \rightarrow \infty} 1/t \log_2 |\lambda_i(t)|, \quad i = 1 \dots, m.$$

The  $Li$  are ordered in descendent fashion with  $L1 \geq L2 \geq \dots \geq Lm$ . Here  $m$  again is the topological dimension of the lowest dimensional manifold that contains the attractor. Haken (1983) has proved that at least one Lyapunov-exponent is zero for each attractor (except for an equilibrium point). It is the one corresponding to the forward direction of the flow. For dissipative systems the sum of all Lyapunov-exponents is less than zero. A system possessing at least one positive Lyapunov-exponent is by definition called chaotic, which makes the largest

Lyapunov-exponent  $L1$  an important classification measure. A positive  $L1$  expresses the sensitive dependence on initial conditions of chaotic dynamics. Hence, the largest Lyapunov-exponent  $L1$  can be utilized as a measure for the "chaoticity" of a system. In fourth-order systems, for example, three qualitatively different possibilities for deterministic chaos exist: ( $L1 > 0, L2 = 0, L3 < 0, L4 < 0$ ), ( $L1 > 0, L2 > 0, L3 = 0, L4 < 0$ ), and ( $L1 > 0, L2 = 0, L3 = 0, L4 < 0$ ). The following table gives an overview, how dissipative dynamical systems can be classified by the nonlinear measures correlation dimension  $D2$  and largest Lyapunov-exponent  $L1$ .

### Methods

Grassberger and Procaccia (1983) first published an algorithm for the estimation of the correlation dimension  $D2$  through calculating the correlation sum  $C(\epsilon)$ . Since the number of operations increases quadratically with the number of data points, the Grassberger-Procaccia-algorithm is enormously time demanding in case of large data sets. Less time consuming than the computation of the complete correlation sum is the estimation of  $D2$  through the "pointwise dimension" (Farmer et al. 1983). Hereby, the distances of the state vectors to only one or some reference vectors  $x_i$  are calculated. The correlation density  $C_D(\epsilon, x_i)$  depending on the embedding dimension  $D$  is defined as

$$C_D(\epsilon, x_i) = \sum_{j=1; j \neq i}^N \Theta\left(\epsilon - \|x_i - x_j\|\right).$$

In order to avoid local or edge effects the reference points should be chosen arbitrarily over the attractor.

$D2$  finally is estimated by plotting  $\log C_D(\epsilon)$  against  $\log \epsilon$ . In case of a deterministic system the slopes of the

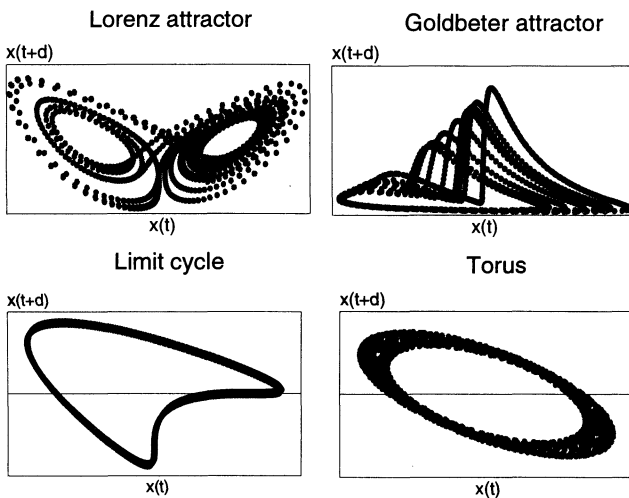


Fig. 2. Examples of two dimensional projections of strange attractors (upper part), a limit cycle, and a torus.

TABLE I

Theoretical values of  $D2$  and  $L1$  for different dissipative dynamical systems

System	Attractor	Correlation dimension $D2$	Largest Lyapunov-exp. $L1$
equilibrium point	point	0	$< 0$
periodic	circle	1	$= 0$
quasiperiodic (k-periodic)	k-torus	k	$= 0$
chaotic	chaotic attractor	noninteger (fractal)	$> 0$

curves  $\log C_D(\epsilon)/\log \epsilon$  are expected to converge against a finite value ("saturate"), if a sufficiently large embedding dimension  $D$  is chosen. Since it is difficult to distinguish between slopes of curves visually, calculating the slopes  $\log C_D(\epsilon)/\log \epsilon$  by linear fits and plotting *versus*  $\log \epsilon$  is more suitable. For random signals (noise) in theory no "saturation" is expected, since stochastic data do not converge to finite dimensional attractors in phase space. Nevertheless, in practice low dimensions can numerically be calculated for stochastic data, as for example Osborne and Provoncale (1989) demonstrated for systems with power-law spectra (a simple class of colored noises). As Grassberger and coworkers (1991) pointed out, these dimension estimates for stochastic time series of finite data length are related to individual trajectories and should not be interpreted as attractor dimensions.

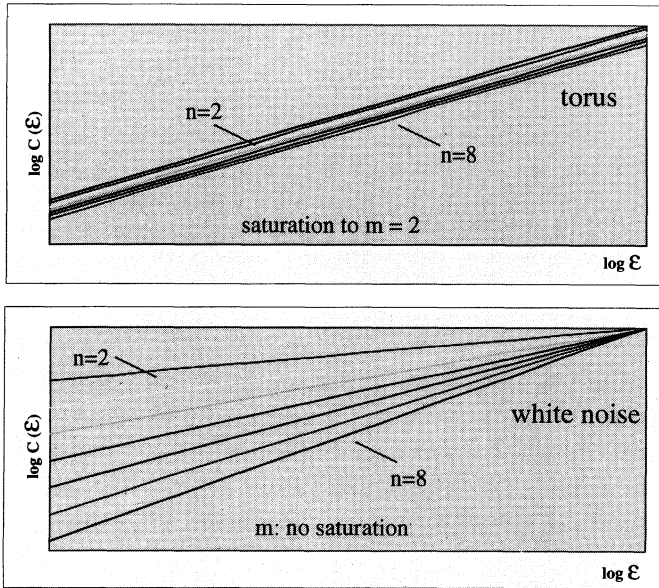


Fig. 3. In case of a two dimensional attractor (torus) the slopes of the curves  $\log C(\epsilon)$  *versus*  $\log \epsilon$  converge to a saturation value of  $m = 2$ . In case of white noise no convergence could be obtained.

Although the computation of correlation dimensions is simple in general, one must pay attention to undesirable effects. One has to be aware that the reconstructed data set is not really equivalent to the system's attractor. Mathematically, attractors are defined by the indecomposable subsets to which trajectories asymptotically converge in phase space for times increasing towards infinity. Thus, an attractor is always a kind of conceptual

object. What we operate on in reality are transient trajectories in phase space for some finite time span. The trajectories themselves have a one-dimensional curve structure, a fact that influences the calculation of  $D_2$ . In order to avoid the curve-structure influence, one should simply leave out the points  $x_i$  for the computation of  $C(\epsilon)$ , which are next in time to the reference points  $x_R$ , that is  $x_i \in [x_{R-w} ; x_{R+w}]$  (Theiler 1986). As correction parameter  $w$  the delay time should be chosen, or at least a value being large enough, so that no alterations in the plots  $\log C(\epsilon)/\log \epsilon$  against  $\log \epsilon$  can be observed for higher values of  $w$ .

A robust routine for the estimation of the largest Lyapunov-exponents was introduced by Wolf et al. (1985). This method is based on the following consideration. An arbitrary vector  $z(0)$  in tangent space of  $x_0$  can be represented by a combination of eigenvectors  $y_k$  of the linearized flow map  $D_{x_0} \Phi_t(x_0)$ :

$$z(0) = \sum a_k y_k$$

After a time  $t$   $z(t)$  can be written as

$$z(t) = \sum a_k m_k(t) y_k$$

where  $m_k$  are the eigenvalues corresponding to  $y_k$ . Equivalently

$$z(t) = m_1(t) \sum a_k m_k(t)/m_1(t) y_k$$

Since the eigenvalues  $m_k(t)$  are ordered descendingly and  $\log |m_k(t)|$  are growing exponentially in time, it is generally the case that

$$\lim_{t \rightarrow \infty} 1/t \log_2 |z(t)|/|z(0)| = \lim_{t \rightarrow \infty} 1/t \log_2 |m_1(t)| = L_1$$

That means that for an arbitrary initial vector the long time evolution generally is determined by the largest Lyapunov-exponent  $L_1$ .

The algorithm starts by computing the distance  $X$  of two nearby points. Because attractors are bounded in practice, the difference vector can not be evolved for infinite times, but has to be renormalized after a certain time span ("evolv"). A new reference point is chosen with the properties to minimize the replacement length and the orientation change. Now, the two points are evolved again and so on. After  $m$  propagation steps, the

first positive Lyapunov-exponent results from the sum over the logarithmic ratios of the vector distances before ( $X$ ) and after propagation ( $X'$ ) divided through the total evolving time:

$$L1 = 1/(t_m - t_0) \sum_{k=1}^m \log_2 [X'(t_k)/X(t_k - 1)]$$

The search for a replacement point extends on a cone centered about the previously evolved vector. Points lying outside the region over which dynamics are assumed to be describable by linearized flow map ("scalmax") and points lying closer than the average noise level ("scalmin") are discarded. If no point can be found, the angle of the cone is expanded and the search goes on. Frank and coworkers (1990) suggested another displacement technique introducing a priority function which depends on the replacement length  $r$  and the orientation change  $\Theta$ :

$$p(r, \Theta) = \{\alpha + \beta[(scalmax - r)/(scalmax - scalmin)]\gamma\} \cdot \cos\Theta$$

In our experience the modified algorithm is less time consuming and converges faster to the final value.

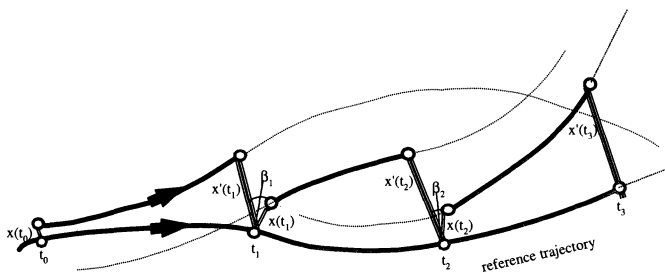


Fig. 4. The principle of calculating the Lyapunov exponent L1. For details see text.

Besides the embedding dimension, the following input parameters are required by the Wolf-algorithm: the minimum scale ("scalmin"), the maximum scale ("scalmax") and the propagation time ("evolv"). Scalmin is given by the expected noise level of the EEG acquisition. Scalmax is the upper boundary for the distance of points considered for replacement. It depends on the maximum amplitude of the signal. According to Wolf and coworkers (1985) one should compute L1 for several values of scalmax to assess the region where the calculations are

approximately stationary. We typically express scalmax in terms of the maximal possible distance ("maxdist") in  $n$ -dimensional phase space corresponding to the maximum amplitude of the EEG epoch under study (e.g. 10% maxdist). Wolf and coworkers suggested that the evolving time should lie between 0.5 and 1.5 of the time needed to transverse the attractor once. In order to estimate the required time the low-frequency part of the power spectra can serve as an approximation. One moreover should be aware, that the time-delay-method introduces a certain structure into the reconstructed data. Suppose the dimension of the embedding space is  $d$ . If two points  $x_i$  and  $x_k$  with distance  $\epsilon$  are propagated about a time equal to the delay time  $\tau$ , they are restricted to a line parallel to the  $d$ -axis. Generally, the maximal possible new distance  $X'_{\max}$  would be  $\sqrt{d} \cdot \text{amp}$ , where amp is the maximum amplitude of the measured data. In case of  $\text{evolv} = \tau$  the maximal possible distance  $X'_{\text{res}}$  is only  $\sqrt{\epsilon^2 + \text{amp}^2}$ . Hence resonance like phenomena occur, if  $\text{evolv} = \tau, 2\tau, \dots, (d-1)\tau$ . In order to avoid these "resonances" the propagation time for each step should be chosen randomly out of a specified interval (Fell and Beckmann 1993). For sleep EEG data for example we typically chose the evolving time randomly out of an interval between 0,2 sec and 0,6 sec corresponding to frequencies between 1,67 Hz and 5 Hz.

White noise in theory yields an infinite largest Lyapunov-exponent, whereas in practice the calculated value will be restricted because of finite data length and the implementation of a minimal distance. For coloured noise data the divergence of subsequent data points may additionally be reduced as a result of low pass filtering, so that a low largest Lyapunov-exponent may be estimated. Therefore, calculation of D2 (as outlined above) and L1 alone can yield no unambiguous information regarding a deterministic or stochastic generation mechanism underlying the data under investigation. So-called surrogate data tests are suitable to provide more evidence. Hereby, control data sets are constructed, which share certain statistical properties with the original data. For example Gaussian-rescaled phase randomized surrogates represent the null hypothesis of linearly autocorrelated noise, but possess the same amplitude distribution and a very similar power spectrum compared to the original data (Theiller et al. 1992, Rapp et al. 1994). Through comparing nonlinear measure estimates for both original and phase randomized data, the linear stochastic model may be falsified. More subtle discriminant measures were proposed by Kaplan and Glass (1992) taking into



account the direction of flow (in practice: trajectories) in phase space, which is the feature being most specific for chaotic systems. However, surrogate data tests only can serve for differentiation of deterministic signals from linear stochastic data, whereas there is still the alternative of a nonlinear stochastic process.

## Findings

Based on the sleep classification according to Rechtschaffen and Kales (1968) sleep stage related changes of nonlinear EEG measures were reported by several groups. It was observed that estimates for D2 and L1 quite robustly decrease with deepness of sleep and that REM sleep yields values in the range of the results for stage I (e.g. Ehlers et al. 1991, Röschke and Aldenhoff 1991, Fell et al. 1993, Pradhan and Sadasivan 1996). Another consistent finding is the decrease of the correlation dimension from eyes opened to eyes closed condition for wake EEG (e.g. Pritchard and Duke 1992, 1997, Elbert et al. 1994). Meyer-Lindenberg (1996) reported a highly significant increase of D2 in the course of normal aging. Alterations of D2 and L1 were moreover described during cognitive tasks, emotional states and under memory load (e.g. Lutzenberger et al. 1993, Sammer 1996, Stam et al. 1996, Aftanas et al. 1997).

The most important application of nonlinear measures to EEG analysis certainly is the detection of pathologically altered mental activity occurring in neurological or psychiatric diseases. The underlying idea is, that nonlinear measures might be more sensitive for pathologically altered mental states than conventional EEG measures are, because of the nonlinear nature of brain dynamics. It was for example detected, that D2 and L1 markedly decrease during and even some time before an epileptic seizure (Iasemidis and Sackellares 1991, Lehnertz et al. 1995, Pijn et al. 1997). Altered nonlinear EEG characteristics were moreover observed in case of schizophrenia for the wake (Elbert et al. 1992, Koukkou et al. 1992) and sleep EEG (Röschke et al. 1993, 1995b), for the sleep EEG in depressive illness (Röschke et al. 1995a), and for the wake EEG in case of Alzheimer's disease (Pritchard et al. 1994, Stam et al. 1996, Besthorn et al. 1997).

In spite of the many reports of significant changes of nonlinear EEG features during different mental states, still the question is crucial, whether conventional stochastic measures are not sufficient to perform the characterization of EEG states. In recent investigations, it

was shown that the discrimination of sleep stages, as well as the differentiation of wake EEG in Alzheimer's disease *versus* healthy controls is indeed clearly improved by addition of nonlinear measures (Pritchard et al. 1994, Fell et al. 1996, Stam et al. 1996). In these studies, best discrimination results were gained by combinations of spectral (stochastic) and nonlinear (deterministic) EEG measures. Thus, the view is supported, that EEG signals in first order can be classified by conventional stochastic measures, but that a complete description has to include nonlinear characteristics.

## UNRESOLVED PROBLEMS IN NONLINEAR EEG ANALYSIS

The essential new element, which is introduced by nonlinear EEG methods compared to conventional analyses, is the language of phase space, meaning a multi-dimensional representation of system states. Extraction of characteristic phase space features through estimation of nonlinear measures can provide information about system dynamics being unaccessible to other methods. The definition of nonlinear measures like D2 and L1 hereby is based on the concept of attractors in phase space. On the one hand the attractor concept incorporates the idea of an autonomous (stationary) system, meaning the dynamical equations do not explicitly depend on time. Analysis of a nonautonomous (nonstationary) system would yield a mixing or a disturbance of dynamical properties. Changes of system dynamics (bifurcations) for example result in jumps in the contributions to the largest Lyapunov-exponent, since connections to initial states are distorted at the change points (see e.g. Kowalik and Elbert 1994, 1995). Hence, in case of nonstationary dynamics it is unclear, whether observed variations in nonlinear measures should be attributed to changes in dynamical properties or just to side effects of nonstationarity. In previous works, we estimated D2 and L1 for different sleep stages classified according to visual scoring criteria (Rechtschaffen and Kales 1968). We found a high intraindividual consistency of nonlinear estimates during a single sleep stage. This led us to the assumption that the EEG during each of the different sleep stages can in first order be regarded as a signal stemming from an autonomous system. Nevertheless, the dynamical microstructure within sleep stages, the course of transitions between stages and a reconsideration of the conventional sleep classification from a dynamical perspective remained open questions.

The attractor concept on the other hand implicates the investigation of long time periods, since attractors are defined by the convergence of trajectories in phase space for times increasing towards infinity. This means, the accuracy of reconstruction in phase space is determined by the number of data points. According to Wolf and Bessoir (1991) one has to pay attention to three relevant quantities. The first is the number of data points per "orbit" necessary to get a minimal sense of orbital continuity. This point is uncritical, if the sampling frequency is chosen adequately in order to catch the prominent frequencies of the power spectrum. The second quantity is the number of orbits required to reproduce phase space structures (for the calculation of D2) and obtain a long time average (for the computation of L1). It is plausible that this number should exponentially grow with  $d-1$ , where  $d$  is the dimension of the submanifold that contains the system's attractor. The third factor was called "density factor" by Wolf and Bessoir (1991) and describes the number of data points required to fill in the local attractor structure in phase space. This number should grow exponentially with  $d$ . If it is too small, fractal scaling will suffer and replacement points for the computation of Lyapunov-exponents will not be close enough. An optimistic estimate for the number of data points needed for the calculation of correlation dimension and largest Lyapunov-exponent would be 10 points per orbit,  $10^{d-1}$  orbits and  $10^d$  total data points. Since dimension and Lyapunov calculations provide no direct feedback on the appropriateness of the data set's size, they can only be taken as estimators for the dynamical properties of the underlying system. The tendency of variations of nonlinear estimates however may be preserved even with short time series, as was demonstrated by Preissl and colleagues (1997) in case of dimension calculations. Since EEG data appear to be high dimensional (dimensions for nonpathological EEG data are in any case above 4) and available data length from a theoretical point of view therefore are too small, the absolute values of nonlinear EEG measures should be regarded with care and interpretations should be based on variations between different conditions respectively mental states. Anyhow, calculations of nonlinear measures only make sense, when at least several orbits in phase space are accessible, which for EEG data means a time window of at least some seconds duration.

From the above said, the dilemma of the nonlinear EEG researcher is obvious. On the one hand, the analyzed EEG time series should be as long as possible in

order to provide a sufficient accuracy of phase space reconstruction. On the other hand the EEG should be stationary, which is more likely in case of short time intervals, since electric brain dynamics exhibit pronounced changes (bifurcations) on a large time scale as for example during night sleep. Unfortunately, concatenation of several short EEG segments, which belong to a certain mental state and therefore are assumed to be stationary, offers no general solution. Time-delay reconstruction can not be done across segment borders, since otherwise data points not belonging together in time would be composed to vectors. Therefore,  $\text{delay} \cdot (\text{dim}-1)$  time points at the end of each subsegment have to be omitted from reconstruction. If the data length of the concatenated EEG epoch is nevertheless sufficient, estimations of correlation dimensions are unaffected by the interruptions in the time series since they rely on metric calculations. However, Lyapunov calculations may yield erroneous results, since the trajectory course is distorted at the edges of the subsegments. In order to obtain reliable estimates for nonlinear EEG measures, it will therefore in general be necessary to select uninterrupted epochs, which are as long as possible, but still guarantee stationarity. Since nonlinear methods themselves are based on the assumption of stationarity, they can *per se* give no reliable information about this aspect. In the last paragraph, we hence propose a combined strategy for the dynamical analysis of EEG signals incorporating both stochastic and nonlinear deterministic methods.

## THE STOCHASTICAL APPROACH: CONCEPTS, METHODS AND FINDINGS

### Concepts

In general EEG signals may be analyzed by any methods capable to catch persistent signal features which are of interest within the considered context. The potentialities of nonlinear methods for EEG analysis were demonstrated in previous paragraph. This approach appears to be the most adequate one in the context of a global description of brain activity as stemming from a nonlinear system of high complexity.

The question may arise, whether data generated by a nonlinear deterministic system could be treated adequately by stochastic methods at all. In answer to this question, one may remark, that stochastic methods are well-established within nonlinear systems theory (e.g.

Parker and Chua 1989). System bifurcations like period-doubling or intermittency for example can be observed through estimation of power spectra. Moreover, the concepts for nonlinear measures like correlation dimension or Kolmogorov-entropy are based on probabilistic descriptions. The definition of Renyi-dimensions (the correlation dimension is a special case of a Renyi-dimension) originally is derived from the probability to find system states within volume elements of an infinitesimal diameter. This definition relies on the fact, that in case of a chaotic system, the states in phase space converge to a fractal attractor set. In this sense probabilistic descriptions are already incorporated in nonlinear systems theory.

In order to illustrate the relation between the trajectory course of a chaotic system and the distribution probabilities in phase space, let us consider a simple example. The following scalar equation (in discrete time) defines the so-called Logistic Map (with the control parameter  $\alpha = 4$ ):

$$x(t+1) = 4x(t)(1 - x(t)), t=0,1,2,\dots, x \in [0,1] \quad (1)$$

It is well known that this system shows chaotic behavior. Pick an initial state  $x(0)$  and calculate a long trajectory:

$x(0), S(x(0)), S^2(x(0)), \dots, S^N(x(0))$  of length  $N$ , where  $S(x) = 4x(1-x)$ .

Then it is straightforward to determine the fraction, call it  $f$ , of the  $N$  system states that is in the  $i$ -th interval from

$$f = k_i / N, i = 1, \dots, n, \text{ with } n \ll N,$$

where  $k_i$  = number of  $S^j(x(0)) \in [(i-1)/n, i/n], j = 1, \dots, N$ , so that  $\sum k_i = N$ .

In such a way, we can construct a histogram of the trajectory. It turns out that the histogram looks like almost the same for trajectories started from any initial condition if  $N$  is sufficiently large. More precisely, it can be proved that as  $N$  tends to infinity the histogram tends to the limit probability density:

$$f^*(x) = \frac{1}{\pi\sqrt{x(1-x)}} \quad (2)$$

So, if we consider any set of initial conditions and will calculate the trajectories started from these initial condi-

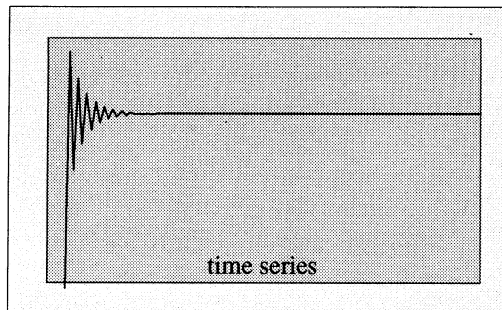
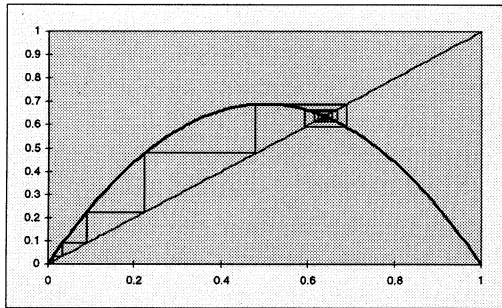
tions then the probability density of these trajectories will tend to the same limit  $f^*(x)$  without regard of initial conditions. Theorems of such type can be proved in much more general situations (Lasota and Mackey 1995). Hence, despite the fact of chaotic behavior of individual trajectories the state distribution in phase space is quite regular (of course under certain conditions which we can not describe here). By this way, the probabilistic description of chaotic systems arises naturally because the chaotic system generates a probability density (it is important to underline that the chaotic system has no random elements, but that the behavior of its output - from the point of view of an external observer - is random). So it is appropriate (and necessary) to use all ordinary probabilistic characteristics (for example, single- and multivariate distributions, correlation functions, spectra and so on) and methods from statistical analysis for the investigation of chaotic systems.

A stochastic process is actually a reasonable model for the evaluation of EEG activity in any case. A sample record of EEG data may be regarded as one single realisation of a random process. However, one must take into consideration that the EEG as a random process is a mathematical abstraction rather than a physical reality (see paragraph 1). It was suggested at the beginning of mathematics based EEG investigations that EEG records belong to the type of random data which can not be described by explicit mathematical equations, but can be described in statistical terms, i.e. by probability distributions, means, variances, covariances, spectra and so on (e.g. John 1977, Dumermuth and Molinari 1987). Autoregressive modeling and spectral analysis are main techniques that consider the EEG as a stochastic process (e.g. Nuwer 1988). Many publications of the last ten years testify that spectral and autoregressive approaches to EEG analysis are fruitful both for clinical applications and for fundamental research (e.g. John et al. 1977, 1988, Duffy 1986). Also period analysis is practiced widely in EEG research (Sharp et al. 1975, Harner 1977, Lorig 1986). This time domain technique quantifies the number of waves occurring in different frequency bands. Spectral and period analyses differ in that spectral analysis decomposes the EEG into sine waves of different frequencies and yields a measure of the power (squared amplitude) of waves in different frequencies. Period analysis indicates the actual number of waves which occur in a particular frequency range (and thus the dominant frequency) by timing successive zero-crossings with respect to baseline.

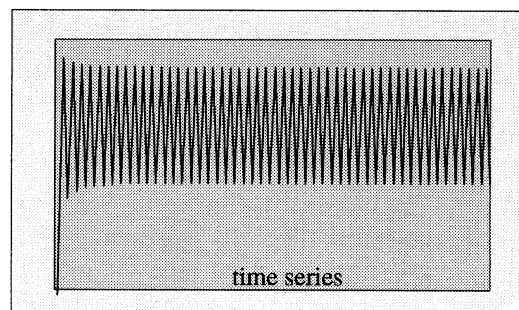
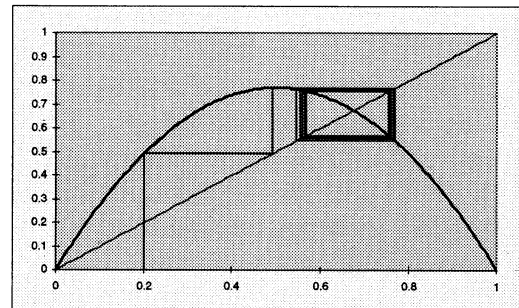
**A**

# LOGISTIC MAP

**FIXPOINT**  
 $\alpha = 2.76$



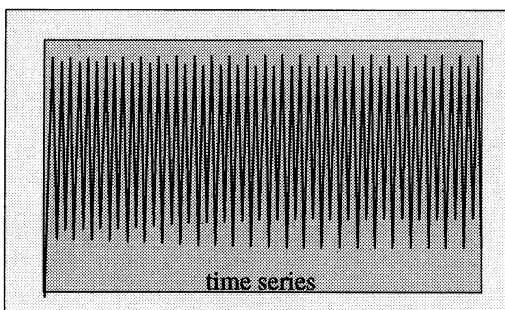
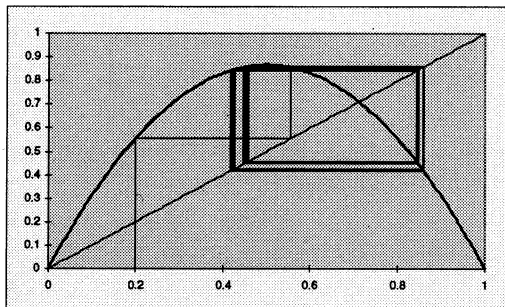
**LIMIT CYCLE**  
 $\alpha = 3.08$



**B**

# LOGISTIC MAP

**PERIOD 4**  
 $\alpha = 3.47$



**CHAOS**  
 $\alpha = 3.72$

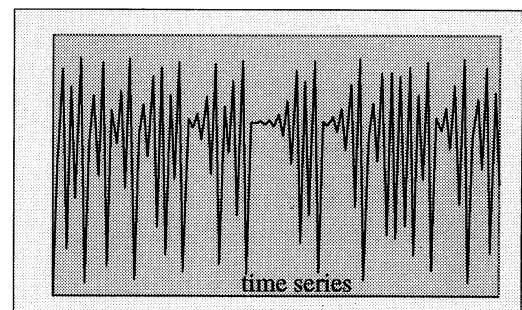
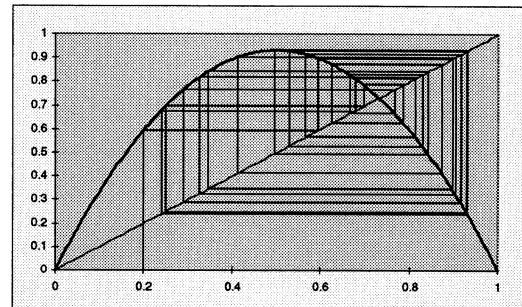


Fig. 5. A, the logistic map comes out with a fixpoint for  $\alpha = 2.76$  and with a limit cycle for  $\alpha = 3.08$ ; B, the logistic map comes out with another limit cycle for  $\alpha = 3.47$  and with deterministic chaos for  $\alpha = 3.72$ .

There are some stochastic methods which use results of autoregressive, spectral or period analysis. One of the main among these methods is principal component analysis (PCA), which is also widely used for the analysis of event related potentials (ERPs). PCA relies on the mathematical technique of singular value decomposition, by which a matrix is decomposed into linear independent basis vectors. This technique is very useful in case one needs to find only a few factors or sources for explanation of the main contributions to EEG variance. With respect to analysis of spontaneous EEG, PCA may be applied to identify basic frequency patterns of the broad band EEG spectrum (Lorig and Schwartz 1989, Makeing and Jung 1995). Other methods implement sets of spectral parameters or autoregressive coefficients as vectors for multidimensional reconstructions. Among these are different types of clustering (e.g. Jansen et al. 1981a,b, Gath et al. 1983), syntactic analysis (Giese et al. 1979), so-called learning vector quantization (Kohonen 1980, Pregenzer et al. 1994) or artificial intelligence methods (e.g. Jagannathanand et al. 1982).

However, there is one crucial problem for any method of EEG analysis, which is the basic nonstationarity of EEG signals (e.g. Lopes da Silva 1975, Jansen 1981). For example, it was reported that the proportion of non-stationary segments in EEG recordings is increased from 10 to 80% when extending a section of EEG analysis from 1 s to 16 s (McEwen and Anderson 1975). It is a common conception that rather short homogenous segments of spontaneous EEG activity represent the basic blocks (Lehmann et al. 1987) or operative states (Basar and Bullock 1992) of information processing. However, all known stochastic methods may be useful only for stationary periods of EEG. The same problem arises for nonlinear analysis, which is based on the assumption that the system under study is an autonomous one. This is the reason to study the EEG by means of so-called segmentation methods, on which we will focus in the following.

### **Parametric methods for the detection of EEG nonstationarity**

Bodenstein and Praetorius (1977) suggested that the EEG can be described as a piecewise stationary process, i.e. that EEG data are "glued" from stationary segments with different probabilistic characteristics. To obtain an adequate description of a piecewise stationary realization, the data set must first be divided into segments with different characteristics by determining the points

of the "adhesion" or "gluing". When this problem is solved, mathematical models can be adjusted for each of the stationary "pieces".

Moments of changes in EEG characteristics can be of interest per se (Skrylev 1984, Deistler et al. 1986, Gath et al. 1991) in particular, as indicators of time moments, in which brain activity is changing (Kaplan et al. 1997). Therefore, it is important to find a way to determine the moments of "adhesion" (moments of changes in characteristics) in the EEG as precisely as possible. In mathematical statistics the problem described above is known as the change-point problem (from here on we will use the term "change-point" for a moment of change in some probabilistic characteristic of a random process). In EEG analysis the term "segmentation" is used, which means dividing a EEG into homogeneous (stationary) segments, although some of the authors dealing with the same problem emphasize its particular aspects by the use of specific terms, e.g., "jumping point" (Deistler et al. 1986). The methodology of EEG segmentation of course can be also applied to magnetoencephalographic (MEG) signals, but also to other biomedical signals (e.g. Moussavi et al. 1996).

The most common approach to EEG segmentation is based on the construction of models for the EEG signal. This approach includes primarily autoregressive modelling and computation of autocorrelation functions (Bodenstein and Praetorius 1977, Michael and Houchin 1979, Sanderson et al. 1980, Bodenstein et al. 1985, Penczek et al. 1987, Aufrichtig et al. 1991; see also for a comprehensive review on segmentation methods Barlow (1985)). Crosscorrelation was used by Gath and coworkers (1991) to provide a "dual channel segmentation". In this method EEG spectra are compared in a reference window and in a window which is moving along the EEG. When the difference between the spectra in the two windows exceeds a certain threshold, a segment boundary is placed. A similar method suggested by Deistler and coworkers (1986) is based on the computation of regression models for alpha power spectra.

Although these methods of EEG segmentation appear to be quite sensitive for some diagnostic goals in clinics (e.g. Creutzfeldt et al. 1985, Merrin et al. 1990, Ihl et al. 1993, Lehmann et al. 1993) they have a serious internal contradiction. Both the regressive models and correlation functions used in these methods require a stationary realization for the correct estimation of parameters. Since the stationary intervals can be determined only as

a result of the analysis, a "vicious circle" emerges. This drawback is aggravated when complex models are used, because such models can be adjusted only for relatively long realizations, which often are not stationary. Simple models, in their turn, can not provide high sensitivity, since a large number of such models can be adjusted for a single realization with nearly the same accuracy of the approximation.

Time-varying autoregressive modeling was proposed to overcome these shortcomings (Amir and Gath 1989). In this method, the autoregressive coefficients are represented as a sum of a finite number of known functions of time. The problem of detection of changes in the autoregressive model parameters is thus replaced by the problem of estimation of the coefficients of this expansion, and the same vicious circle appears at higher level of description. A combinatory approach for the EEG segmentation was described by Reschenhofer and co-workers (1987). This approach transforms the problem of EEG segmentation to a complicated optimization problem. The proposed algorithm however seems to be able to detect in practice only the strongest change-points. Another combinatory approach (Bayesian approach) based on the maximum likelihood principle was suggested by Biscay and coworkers (1995), but knowledge of the change-point apriori distribution is indispensable for solving this extremal problem, which is evidently unknown in most of practical situations.

A different way of segmentation is based on dividing the EEG into sequential epochs, generally with fixed length, followed by classification of these epochs. The boundaries are placed between epochs, which were found to belong to different classes (Pascual-Marqui et al. 1995). It is evident that these boundaries, in general, are not equivalent to the statistical concept of change-points. The time resolution of these methods is limited by the epoch length. Other methods of EEG segmentation described in literature (e.g. Skrylev 1984, Lehmann et al. 1987) also do not use the methods of mathematical statistics for determining change-point moments, and hereby can provide only a rough estimation of the EEG segmentary structure. It can be concluded that all known parametric methods of EEG segmentation have features, which essentially limit their efficiency and validity. An additional point to emphasize is, that most of the EEG segmentation methods which have been reported previously were designed for the analysis of a single or only a few EEG characteristics (e.g., the spectrum), whereas it may be

desirable to be able to utilize the numerous characteristics used for the description of the EEG.

In order to overcome the problems discussed above, we designed a method for EEG segmentation based on a nonparametric approach to the change-point detection problem (Darkhovsky 1976, 1984, Brodsky and Darkhovsky 1989, 1993a,b). This method provides statistically justified segmentation (with estimation of confidence intervals for change points) for practically all types of the EEG characteristics. The advantage of the nonparametric methods is that they minimize the need for apriori information about the analyzed signal, which is difficult or impossible to obtain in the case of EEG analysis.

### **Nonparametric approach to the detection of EEG nonstationarity**

Before describing the main ideas of the nonparametric approach to the problem of detecting changes of probabilistic characteristics, let us explain how these changes can be defined formally. We suppose that an observed random process is "glued" from several strictly stationary processes. It is well-known that the full description of a random process, in general, is represented by the total set of its finite-dimensional distributions, which are invariant with respect to time shifts for strictly stationary processes. Suppose that stationary processes - components of a glued process - differ in their distribution function. The points of "gluing" then are called the moments of changes of statistical characteristics. For a strictly stationary process any statistical characteristic does not depend on time, and therefore moments of changes in the piecewise stationary scheme are determined by moments of gluing. If the description of a nonstationary process is known up to some parameters, and these parameters, in their turn, are piecewise stationary (in the simplest case - piecewise constant functions of time), then moments of changes in characteristics of these parameters are change-points of the original process.

Our methodology is based on two main ideas. The first idea is the following. It can be proven (Brodsky and Darkhovsky 1993b), that the detection of changes in any distribution function or some other probabilistic characteristic can be reduced (with any degree of accuracy) to the detection of changes in the mathematical expectation of some other random sequence formed by the initial one. For example, if the autocorrelation function of a sequence changes, then considering new sequences

$$V_t(\tau) = x_t x_{t+\tau}, \tau = 0, 1, 2, \dots,$$

we will reduce the problem to detection of changes in one of the sequences  $V_t(\tau)$ . The sequence  $V_t(\tau)$  hereby consists of the autocorrelation values  $V_t$  for a specific delay  $\tau$ . This circumstance enables us to limit ourselves to the development of only one basic algorithm for detection of changes in the mathematical expectation, and not to create an infinite family of algorithms for detection of changes in arbitrary statistical characteristics. A new sequence constructed from the initial one, in which a change in expectation occurs, will be called a diagnostic sequence.

The second idea of our approach is to detect change-points using the following family of statistics:

$$Y_N(n, \delta) = \left[ \left(1 - \frac{n}{N}\right) \frac{n}{N} \right]^\delta \left[ \frac{1}{n} \sum_{k=1}^n x_k^N - \frac{1}{N-n} \sum_{k=n+1}^N x_k^N \right], \quad (3)$$

where  $0 \leq \delta \leq 1$ ,  $1 \leq n \leq N-1$ ,  $X = \{x_k^N\}_{k=1}^N$  is the realisation under investigation. This family of statistics (3) is a generalized variant of the Kolmogorov-Smirnov statistic, which is used for testing coincidence or difference of distribution functions of two samples (with fixed  $n$ ). In simple words, we calculate the difference between an arithmetic mean of the first  $n$  samples and an arithmetic mean of the last  $N-n$  samples modulo a factor depending on  $\delta$ . This calculation has to be done for all  $n$ ,  $1 \leq n \leq N$ . Then, we compare the maximum of the differences over  $n$  ( $1 \leq n \leq N$ ) with a special threshold. The threshold is calculated on the base of the limit (under  $N$  tends to infinity) characteristics of the statistic. We make a decision about stationarity of the EEG realization, if this threshold is not exceeded, whereas in the opposite case we detect a change-point.

It can be shown, that the above defined family of statistics (3) gives asymptotically (as  $N$  tends to infinity) optimum estimates for the change-points under weak mathematical assumptions (Brodsky and Darkhovsky (1993b)). An important property of these statistics is that the choice of  $\delta = 0$  provides the minimum for the false alarm probability (i.e., the probability for a decision about the presence of change-points, when no change occurred). On the other hand,  $\delta = 1$  corresponds to the minimum probability of false tranquility (i.e., the probability for the absence of change-points, when there actually was a change) and the choice  $\delta = 1/2$  guarantees a minimal estimation error (in time) for a change-point.

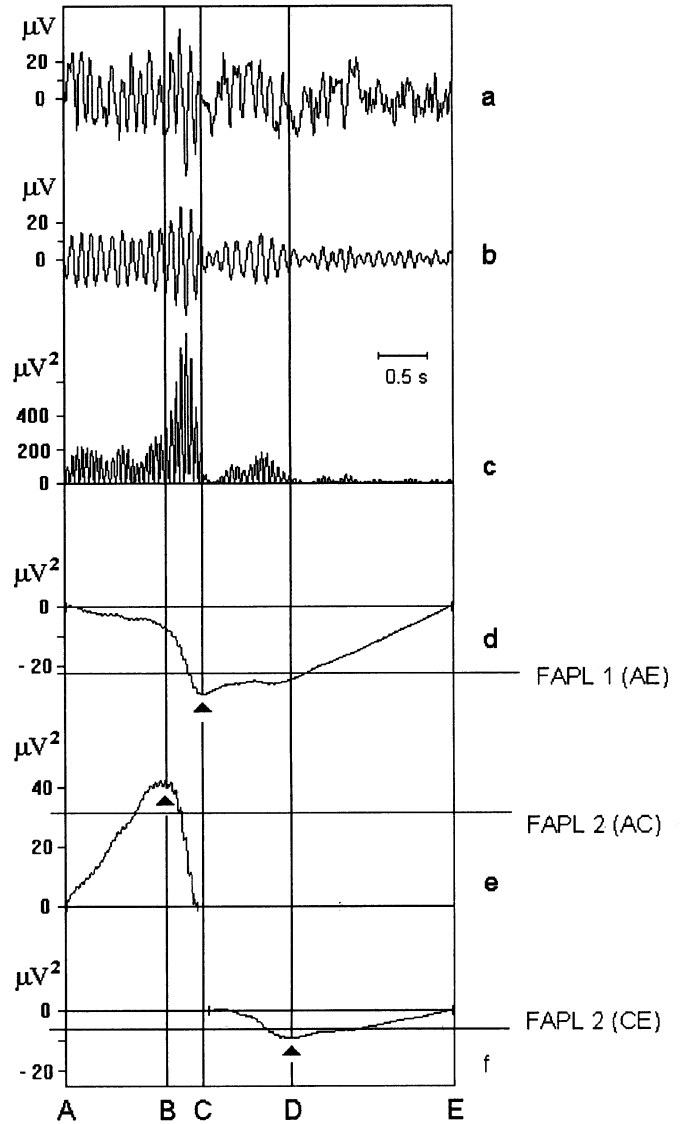


Fig. 6. Principles of segmentation: to obtain the diagnostic sequence, the EEG (a) was filtered in the alpha range (bandpass 7.5-12.5 Hz) (b) and then the consecutive amplitudes were squared, i.e.  $V_t(0)$  was calculated (c). Statistics  $Y_N(n, 1)$  was first computed for the entire length of the diagnostic sequence AE (d). The maximum when exceeding the threshold gave the estimate of the first change-point moment C. For the subsamples AC and CE another statistics were computed ((e) and (f)), and new estimates of the change-points moments B and D were detected. In all cases the thresholds were calculated corresponding to False Alarm Probability Level (FAPL) 0.05. FAPL 1 and FAPL 2 mean first and second stages of the algorithm. Then statistics were computed for the subintervals AB, BC, CD, DE (not shown); since they did not exceed the threshold, the preliminary estimation was terminated, resulting in a set of change-point estimates B, C and D. Triangles indicate maxima of the statistics; solid vertical lines show the moments of preliminary change-point estimates.



The basic parameter, which has to be specified by the user for the threshold calculation, is the false alarm probability. The lower the false alarm probability, the larger is the threshold for change point detection, and the larger are the changes in the characteristic under consideration, that will be detected. By adjusting the false alarm probability it is therefore possible to either focus on the analysis of macrostructural changes (high threshold analysis) or to investigate the microstructural organisation of EEG (low threshold analysis). The main processing steps of the nonparametric change-point method adapted for EEG analysis (Kaplan et al. 1997a, Shishkin et al. 1997) are listed in the following.

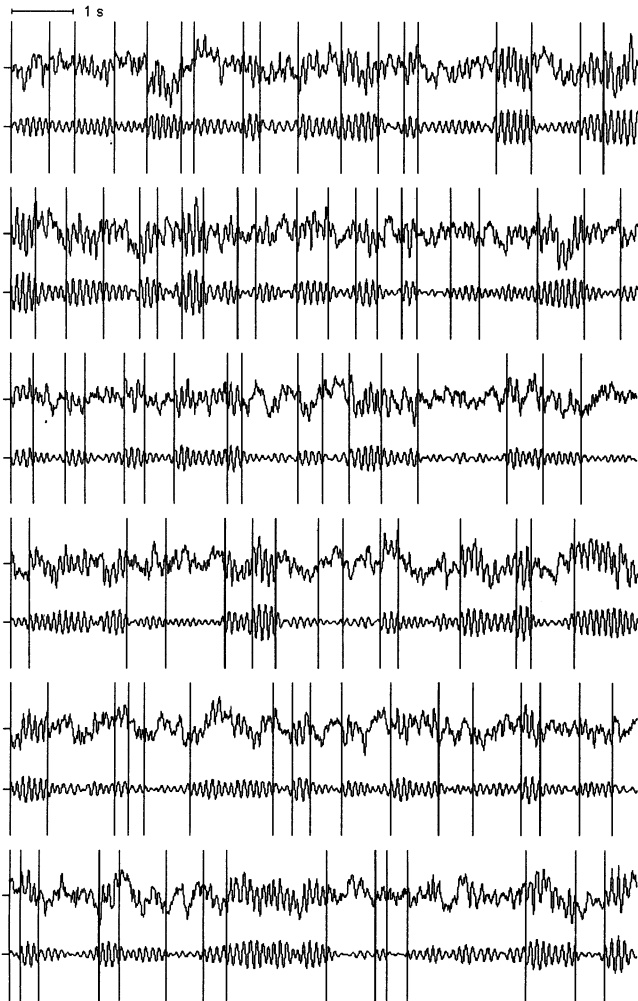


Fig. 7. Segmentation of a continuous EEG recording of 1 min recorded from O2, eyes closed. In each pair of the curves the upper part is EEG signal, and the lower is the same EEG after filtering in alpha band (7.5-12.5 Hz). Vertical lines show the change-points in spectral power of the alpha band.

## Nonparametric change point analysis of EEG data: processing steps

### CALCULATION OF THE DIAGNOSTIC SEQUENCE

We typically construct the diagnostic sequence from autocorrelation values derived from the original EEG data:

$$V_t(\tau) = x_t x_{t+\tau} \quad \tau = 0, 1, 2, \dots,$$

Changes in autocorrelation values correspond to variations in power spectra, since the power spectrum is equal to the Fourier transform of the autocorrelation function. In particular, the mean of  $V_t(0)$  is identical with the total power (Parseval's theorem).

### CHECKING THE HOMOGENEITY HYPOTHESIS

Compute the value  $\max \{|Y_N(n, \delta = 1)|: 1 \leq n \leq N - 1\} = \eta_N$  and the threshold  $C$  (note that  $\delta = 1$ , i.e. the probability of false tranquility is minimal). The threshold is computed on the basis of the limit theorem in dependence on the given false alarm probability, which is set rather high at this stage. If  $\eta_N \leq C$ , then the homogeneity hypothesis is accepted (i.e., the absence of disorders) and the procedure is completed; in the other case, we go to the next step.

### PRELIMINARY ESTIMATION OF CHANGE-POINTS

The global maximum of the statistic  $|Y_N(n, \delta = 1)|$  call it  $n_1$ , is assumed to be the estimate of the first found change-point. Now, two new samples:

$$Z_1: 1 \leq n \leq n_1 - [\epsilon N] \text{ and } Z_2: n_1 + [\epsilon N] \leq n \leq N$$

are formed. Here  $\epsilon$  is a number, which is computed by the volume of the sample and the steepness of the statistics' maximum and gives the preliminary estimate of a confidence interval for the change-point. Then each of the new samples  $Z_1$  and  $Z_2$  is checked for homogeneity (step 1), and if not the case, we go again to step 2. The procedure is repeated until we obtain statistically homogenous segments. As a result of step 2 we obtain a set of preliminary estimates of change-points  $n_1, n_2, \dots, n_k$ , where  $k$  is the preliminary estimate of the number of change-points.



### REJECTING OF DOUBTFUL CHANGE-POINTS

The following subsamples are formed ( $s=2, \dots, k-1$ ):

$$X_1 : 1 \leq n \leq n_1 + \frac{1}{2} (n_2 - n_1)$$

$$X_s : n_{s-1} + \frac{1}{2} (n_s - n_{s-1}) \leq n \leq n_s + \frac{1}{2} (n_{s+1} - n_s)$$

$$X_k : n_{k-1} + \frac{1}{2} (n_k - n_{k-1}) \leq n \leq N$$

Thus, inside each subsample  $X_i$  there is a single preliminary change-point estimate  $n_i$ . Each sample is analyzed in analogy to stage 2, but with a lower false alarm probability. If the homogeneity hypothesis is accepted for a sample, then the corresponding change-point is rejected.

### FINAL ESTIMATION OF CHANGE-POINTS

For each sample  $X_i$  (of the volume  $N_i$ ) remaining after step 3 the statistic  $Y_{N_i}(n, \delta = 0)$  is computed. The maximum point of the module for this statistic is assumed to be the final estimate of the  $i$ -th change-point. Then the confidence interval is computed from the statistic  $Y_{N_i}(n, \delta = 1/2)$ .

### INTEGRATION OF RESULTS FOR DIFFERENT DIAGNOSTIC SEQUENCES AND MULTIPLE ELECTRODES

Only change points, which simultaneously (within the confidence intervals) occur for all diagnostic sequences under consideration are finally chosen, in the other case points are rejected. Data stemming from several EEG channels may be integrated in the same way. Topographically adjacent channels may be combined to functional groups, and only changes simultaneously occurring in all channels belonging to one functional group may be accepted.

### Findings

Our previous results from the application of the non-parametric change point method to EEG analysis are related to the detection of microstructural changes in EEG records. For this purpose, we implemented a rather high false alarm probability of 5%, which results in a low change point detection threshold. It turned out, that the distribution of EEG change points can be generally also

recognized from visual inspection in case of alpha band filtered data, but not in case of broad-band EEG (Shishkin et al. 1997). For multichannel records of spontaneous wake EEG, we found a characteristic topographic structure of change-point coincidences, which depends on the cognitive state under investigation (Kaplan et al. 1997a).

Based on our approach we also calculated change-points for different stages of sleep EEG for healthy subjects (Kaplan et al. 1997b). For change-point processing we previously extracted from a sleep EEG record segments defined as REM and NREM (stages 2 and 4 separately) on the base of expert scoring according to Rechtschaffen/Kales criteria. Each segment had 16,384 data points and for every subject we extracted 5 realizations of each stage without any artifacts. We subsequently calculated the number of change-points within the given stages and the total distribution of segment lengths between change-points. The mean time interval between change points for the low threshold analysis was approximately 3-9 s. Although the distribution of change-points showed individual features across subjects, there was a strong difference for each subject between stages 2 and 4. The sleep stage REM was characterized by an intermitted position. For example, for a single subject 80% of segments between change-points (SBCP) in the EEG were no longer than 1,200 ms for sleep stage 4 and no longer than 1,550 ms for sleep stage 2. Over seven healthy subjects SBCP for stage 4 decreased to 68-85% compared with stage 2. These changes of the SPCP were statistically significant. Our previous findings hence indicate, that sleep stages can be distinguished in their microstructural organisation by means of low threshold change point analysis. In the last paragraph, we will propose a combined nonlinear deterministic and stochastic strategy. This strategy incorporates the macrostructural segmentation of EEG signals through implementation of high threshold change point analysis.

### A COMBINED STRATEGY

As described above, the application of nonlinear methods is based on the assumption of autonomous (stationary) dynamics. On the other hand, it is clear that electric brain activity can not be described by a stationary model, since pronounced alterations in EEG dynamics occur during sleep or during different mental awake states. Hence, a quasistationary description appears

more adequate, namely that nonlinear brain dynamics are piecewise stationary, i.e. that autonomous dynamics are prevalent for finite time intervals and that dynamical bifurcations occur between these intervals. In the following, we will concentrate on the analysis of sleep EEG, which is known to exhibit pronounced dynamical changes, that can be recognized already by visual inspection.

In previous investigations of nonlinear EEG dynamics during sleep, we regarded EEG records corresponding to one of the sleep stages, which were classified according to visual scoring criteria (Rechtschaffen and Kales 1968) as signals generated by autonomous dynamics. Indeed, the variation of D2 and L1 during each of the sleep stages turned out to be small compared to the overall variations between stages (e.g. Röschke et al. 1995a). However, analysis of power spectra for example revealed, that changes in EEG spectra during the night are more or less continuous (Mann et al. 1993), a finding, which is incompatible with the picture of homogeneous and sharply separated states. Moreover, it appears questionable, whether the conventional classification of sleep dynamics into 5 macrostructurally different stages is dynamically justified. Stages 3 and 4 for example (the slow wave stages) are not differentiated *via* qualitative EEG characteristics, but are distinguished quantitatively by their amount of delta waves.

In order to pursue a strategy, which is based on objective dynamical criteria, we started to implement a combined nonlinear deterministic and stochastic approach to EEG analysis. The first step consists of macrostructural segmentation of spontaneous EEG by means of nonparametric change point analysis. For this purpose, we implement a low false alarm probability (as explained in the last paragraph), which corresponds to a high detection threshold. For the investigation of sleep EEG, this threshold should guarantee, that compared to conventional sleep scoring at least roughly the same number of changes can be detected. However, change points should be allowed to occur more frequently and the acceptance of change points should in no way depend on the conventional sleep classification. In this way, the notion of several macrostructurally different sleep stages is basically preserved, but classical sleep analysis will be revised.

After change point segmentation, nonlinear (D2 and L1) and spectral measures (e.g. band powers, spectral coherence) can be estimated for the quasistationary epochs. In case of the nonlinear estimates, a fixed epoch length being as large as possible (e.g. in the order of 1 min) has to be specified (see paragraph 3) and segments

with a smaller duration have to be omitted from analysis. The information from nonlinear and spectral measures may be subsequently integrated by a cluster analysis. Hereby, the question can be addressed, in how far the conventional sleep classification scheme is reflected by the dynamical measures, respectively whether there is evidence for a different macrostructural organisation of sleep EEG. Moreover, since dynamical variations within stages are expected to be larger for the segmentation based approach, the question of homogeneity of sleep stages can be subjected to critical reanalysis. As a last point, surrogate data tests will profit from the combined approach, since also this method is based on the assumption of a stationary EEG epoch under consideration (see paragraph 2). Hence, the information will be valueable, whether surrogate data tests on change point segmented quasistationary EEG epochs yield differences to linear stochastic control data, and whether these differences are more or are less pronounced than those found in previous non segmentation-based analyses.

## REFERENCES

- Achermann P., Hartmann R., Gunzinger A., Guggenbuhl W., Borbély A.A. (1994) All night sleep EEG and artificial stochastic control signals have similar correlation dimensions. *Electroencephalogr. Clin. Neurophysiol.* 90: 384-387.
- Aftanas L.I., Lotova N.V., Koshkarov V.I., Pokrovskaja V.L., Popov S.A., Makhnev V.P. (1997) Non-linear analysis of emotion EEG: calculation of Kolmogorof entropy and the principal Lyapunov exponent. *Neurosci. Lett.* 226: 13-16.
- Aihara K., Matsumoto G. (1986) Forced oscillations and bifurcations in squid giant axons. In: *Chaos. Nonlinear science: theory and applications* (Ed. A.V. Holden). Manchester Univ. Press, Manchester, p. 257-269.
- Aihara K., Matsumoto G. (1987) Forced oscillations and routes to chaos in the Hodgkin-Huxley axons and squid giant axons. In: *Chaos in biological systems* (Eds. H. Degn, A. Holden and L.F. Olsen). Plenum, New York, p. 121-131.
- Amir N., Gath I. (1989) Segmentation of EEG during sleep using time-varying autoregressive modeling. *Biol. Cybern.* 61: 447-455.
- Aufrichtig R., Pedersen S.B., Jennum P. (1991) Adaptive segmentation of EEG signals. *Ann. Int. Conf. of the IEEE Eng. Med. Biol. Soc. Proc. Biol. Signals* 13: 453.
- Barlow J.S. (1985) Methods of analysis of nonstationary EEGs, with emphasis on segmentation techniques: a comparative review. *J. Clin. Neurophysiol.* 2: 267-304.
- Basar E. (Ed.) (1990) *Chaos in brain function*. Springer, Berlin.

- Basar E., Bullock T.H. (1992) Induced rhythms in the brain. Birkhäuser, Boston.
- Besthorn C., Zerfass R., Geiger-Kabisch C., Sattel H., Daniel S., Schreiter-Gasser U., Förstl H. (1997) Discrimination of Alzheimers disease and normal aging by EEG data. *Electroencephalogr. Clin. Neurophysiol.* 103: 241-248.
- Biscay R., Lavielle M., Gonzalez A., Clark I., Valdes P. (1995) Maximum a posteriori estimation of change points in the EEG. *Int. J. Bio-Med. Comput.* 38: 189-196.
- Bodenstein G., Praetorius H.M. (1977) Feature extraction from the electroencephalogram by adaptive segmentation. *Proc. IEEE* 65: 642-652.
- Bodenstein G., Schneider W., Van der Malsburg C. (1985) Computerized EEG pattern- classification by adaptive segmentation and probability density function classification. Description of the method. *Comput. Biol. Med.* 15: 297-313.
- Brodsky B.E., Darkhovsky B.S. (1989) Nonparametric methods of detection of the switching instants of two random sequences. *Automat. Rem. Contr.* 50: 1356-1363.
- Brodsky B.E., Darkhovsky B.S. (1993a) An algorithm of aposteriori detection of repeated disorders of random sequence (in Russian). *Avtomatika i Telemekhanika* 1: 62-67.
- Brodsky B.E., Darkhovsky B.S. (1993b) Nonparametric methods in change-point problems. Kluwer Acad. Publ., Dordrecht (the Netherlands).
- Chaitin G. (1975) Randomness and mathematical proof. *Sci. Am.* 232:47-52.
- Christini D.J., Collins J.J. (1995) Controlling nonchaotic neuronal noise using chaos control techniques. *Phys. Rev. Lett.* 75: 2782-2785.
- Creutzfeld O.D. (1983) *Cortex cerebri*. Springer, Berlin.
- Creutzfeldt O.D., Bodenstein G., Barlow J.S. (1985) Computerized EEG pattern classification by adaptive segmentation and probability density function classification. Clinical evaluation. *Electroencephalogr. Clin. Neurophysiol.* 60: 373-393.
- Darkhovsky B.S. (1976) A nonparametric method for detection of the "disorder" time of a sequence of independent random variables. *Theor. Prob. Appl.* 21:178-183.
- Darkhovsky B.S. (1984) On two estimation problems for times of change of the probabilistic characteristics of a random sequence. *Theor. Prob. Appl.* 29: 478-488.
- Deistler M., Prohaska O., Reschenhofer E., Vollmer R. (1986) Procedure for identification of different stages of EEG background activity and its application to the detection of drug effects. *Electroencephalogr. Clin. Neurophysiol.* 64: 294-300.
- Duffy F.N. (1986) Topographic mapping of electrical brain activity. Bitterworths, Boston.
- Dumermuth H.G., Molinari L. (1987) Spectral analysis of the EEG. *Neuropsychobiology* 17: 85- 99
- Elbert T., Lutzenberger W., Rockstroh B., Berg P., Cohen R. (1992) Physical aspects of the EEG in schizophrenics. *Biol. Psychiat.* 32: 595-606.
- Elbert T., Ray W.J., Kowalik Z.J., Skinner J.E., Graf K.E., Birbaumer N. (1994) Chaos and Physiology: deterministic chaos in excitable cell assemblies. *Physiol. Rev.* 74:1-47.
- Ehlers C.L., Havstad J.W., Garfinkel A., Kupfer D.J. (1991) Nonlinear analysis of EEG sleep states. *Neuropsychopharmacology* 5: 167-176.
- Farmer J.D., Ott E., Yorke J.A. (1983) Dimension of chaotic attractors. *Physica D*, 7: 153-180.
- Feigenbaum M.J. (1978) Quantitative universality for a class of nonlinear transformations. *J. Stat. Phys.* 19: 25-52.
- Fell J., Beckmann P. (1994) Resonance like phenomena in Lyapunov-calculations from data reconstructed by the time delay method. *Phys. Lett. A* 190: 172-176.
- Fell J., Röschke J., Beckmann P. (1993). Deterministic chaos and the first positive Lyapunov- exponent: a nonlinear analysis of the human electroencephalogram during sleep. *Biol. Cybern.* 69: 139-146.
- Fell J., Röschke J., Mann K., Schäffner C. (1996) Discrimination of sleep stages: a comparison between spectral and nonlinear EEG measures. *Electroencephalogr. Clin. Neurophysiol.* 98: 401- 410.
- Fell J., Röschke J., Schäffner C. (1996) Surrogate data analysis of sleep Electroencephalograms reveals evidence for nonlinearity. *Biol. Cybern.* 75: 85-92.
- Felleman D.J., Van Essen D.C. (1991) Distributed hierarchical processing in the primate cerebral cortex.. *Cereb. Cortex* 1: 1-47.
- Frank G.W., Lookman T., Nerenberg M.A.H., Essex C., Lemieux J., Blume W. (1990). Chaotic time series analyses of epileptic seizures. *Physica D* 46: 427-438.
- Gath I., Lehmann D., Bar-On E. (1983) Fuzzy clustering of the EEG signal and vigilance performance. *Int. J. Neurosci.* 20: 303-312.
- Gath I., Michaeli A., Feuerstein C. (1991) A model for dual channel segmentation of the EEG signal. *Biol. Cybern.* 64: 225-230.
- Giese D.A., Bourne J.R., Ward J.W. (1979) Syntactic analysis of the electroencephalogram. *IEEE Trans. Syst. Man. Cybern.* 9: 429-435.
- Grassberger P., Proccacia I. (1983). Characterization of strange attractors. *Phys. Rev. Lett.* 50: 346-349.
- Grassberger P., Schreiber T., Schaffrath C. (1991). Nonlinear time sequence analysis. *Int. J. Bifurcation Chaos* 1: 521-547.
- Gray C.M., König P., Engel A.K., Singer W. (1989) Oscillatory responses in cat visual cortex exhibit intercolumnar synchronisation which reflects global stimulus properties. *Nature* 338: 334- 337.
- Haken H. (1983) At least one Lyapunov-exponent vanishes, if the trajectory of an attractor does not contain a fixed point. *Phys. Lett. A* 94: 71-72.

- Harii R., Lounasmaa O.V. (1989) Recording and interpretation of cerebral magnetic fields. *Science* 244: 432-436.
- Harner R. (1977) EEG analysis in the time domain. In: *EEG informatics: a didactic review of methods and applications of EEG data processing* (Ed. A. Remond) Elsevier, Amsterdam.
- Hayashi H., Ishiyuka M., Hiarawaka K. (1983) Transitions to chaos via intermittency in the onchidium pacemaker neuron. *Phys. Lett.* 98: 474-476.
- Hayashi H., Ishiyuka M., Ohta M., Hirakawa K. (1982) Chaotic behavior in the onchidium giant neuron under sinusoidal stimulation. *Phys. Lett.* 88: 435-438.
- Heagy J.F., Hammel S.M. (1994) The birth of strange non-chaotic attractors. *Physica D* 70: 140-153.
- Hodgkin A., Huxley A.F. (1952) A quantitative description of membrane current and application to conduction and excitation in nerve. *J. Physiol. Lond.* 117: 500-544.
- Holden A.V., Winlow W., Haydon P.G. (1982) The induction of periodic and chaotic activity in a molluscan *Lymnaea stagnalis* neuron. *Biol. Cybern.* 43: 169-174.
- Iasemidis L.D., Sackellares J.C. (1991). The evolution with time of the spatial distribution of the largest Lyapunov-exponent on the human epileptic cortex. In: *Measuring chaos in the human brain* (Eds. W.S. Pritchard and D.W. Duke). World Scientific, Singapore, p. 49-82.
- Ihl R., Dierks T., Froelich L., Martin E.M., Maurer K. (1993) Segmentation of the spontaneous EEG in dementia of the Alzheimer type. *Neuropsychobiology* 27: 231-236.
- Jagannathan V., Bourne J.R., Jansen B.H., Ward J.W. (1982) Artificial intelligence methods in quantitative electroencephalogram analysis *Comp. Prog. Biomed.* 15: 249-258.
- Jansen B.H. (1986) "Is it?" and "So what?" - a critical view of EEG chaos. In: *Measuring chaos in the human brain* (Eds. W.S. Pritchard and D.W. Duke). World Scientific, Singapore, p. 83-96.
- Jansen B.H., Hasman A., Lenten R. (1981a) Piece-wise analysis of EEG using AR-modeling and clustering. *Comput. Biomed. Res.* 14: 168-178.
- Jansen B.H., Hasman A., Lenten R. (1981b) Piece-wise EEG analysis: an objective evaluation. *Int. J. Bio-Med. Comput.* 12: 17-27.
- John E.R. (1977) *Neurometrics: clinical applications of quantitative electrophysiology*. Wiley, New York.
- John E.R., Pritchep L., Easton P. (1987) Normative data banks and Neurometrics. Basic concepts, methods and results of norm constructions - In: *Handbook of electroencephalography and clinical neurophysiology* (Eds. P.S. Gevins and A. Remond). Elsevier, Amsterdam, p. 449-495.
- Kaplan A.Y., Darkhovsky B.S., Fingelkurts A.A., Fingelkurts A.A. (1997a) Topological mapping of the sharp transformation synchronization EEG (in Russian). *J. Hig. Nerv. Act.* 47: 32-37.
- Kaplan A.Y., Darkhovsky B.S., Roeschke J. (1997b) Microstructure of the sleep stages: state-shift analysis of the EEG in humans. In: *Proceeding of XXXIII Int. Congress of Physiological Sciences St. Peterburg* 5: P104.11
- Kaplan D.T., Glass L. (1992) Direct test for determinism in a time series. *Phys. Rev. Lett.* 68:427-430.
- Koester J. (1991) Voltage-gated ion channels and the generation of the action potential. In: *Principles of neural science* (Eds. E.R. Kandel, J.H. Schwartz and T.M. Jessel). Prentice-Hall International. 3 Edition, p. 104-118.
- Kohonen T. (1990) The self-organizing map. *Proc. IEEE* 78: 1464-1480.
- Kookkou M., Lehmann D., Wackermann J., Dvorak I., Henggeler B. (1992) Dimensional complexity of EEG brain mechanisms in schizophrenia. *Biol. Psychiat.* 33: 397-407.
- Kowalik Z.J., Elbert T. (1994) Changes of chaoticness in spontaneous EEG/MEG. *Integr. Physiol. Behav. Sci.* 29: 270-282.
- Kowalik Z.J., Elbert T. (1995) A practical method for the measurement of the chaoticity of electric and magnetic brain activity. *Int. J. Bifurcation Chaos* 5: 475-490.
- Lasota A., Mackey M.C. (1993) *Chaos, fractals, and noise. Stochastic aspects of dynamics*. Second edition. Springer-Verlag, Applied Mathematical Sciences, v. 97.
- Lauterborn W., Holzfuss J. (1991) Acoustic chaos. *Int. J. Bifurcation Chaos* 1: 13-26.
- Lehmann D., Ozaki H., Pal I. (1987) EEG alpha map series: brain microstates by space-oriented adaptive segmentation. *Electroencephalogr. Clin. Neurophysiol.* 67: 271-288.
- Lehmann D., Wackermann J., Michel C.M., Koenig T. (1993) Space-oriented EEG segmentation reveals changes in brain electric field maps under the influence of a nootropic drug. *Psychiat. Res.* 50: 275-282.
- Lehnertz K., Elger C.E. (1995) Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neuronal complexity loss. *Electroencephalogr. Clin. Neurophysiol.* 95: 108-117.
- Lopes da Silva F.H. (1975). Analysis of nonstationary electroencephalograms. In: *Quantitative analysis of the EEG* (Eds. M. Matejcek and G.K. Schenk). AEG Telefunken, Konstanz, p. 421-433.
- Lorig T.S. (1986) EEG and task performance: a comparison of three analytic techniques. *Physiol. Psychol.* 14: 130-132.
- Lorig T.S., Schwartz GE (1989). Factor analysis of the EEG indicates inconsistencies in traditional frequency bands. *J. Psychophysiol.* 3: 369-375.
- Lutzenberger W., Elbert T., Ray W.J., Birbaumer N. (1993) The scalp distribution of the fractal dimension of the EEG and its variation with mental tasks. *Brain Topogr.* 5: 27-34.
- Makeig S., Jung T.P. (1995) Changes on alertness are a principal component of variance in the EEG spectrum. *Neuroreport* 7: 213-216.

- Mann K., Röschke J., Bäcker P. (1993) Investigations of the dynamical properties of the sleep EEG in different frequency bands. *Int. J. Neurosci.* 73: 161-169.
- Manneville P., Pomeau Y. (1979) Intermittency and the Lorenz model. *Phys. Lett. A* 75: 1.
- Matsumoto G., Aihara K., Hanyu Y., Takahashi N., Yoshizawa S., Nagumo J. (1987) Chaos and phase-locking in normal squid axons. *Phys. Lett.* 123: 162-166.
- Mc Ewen J.A., Anderson G.B. (1975) Modeling the stationarity and Gaussianity of spontaneous electroencephalographic activity. *IEEE T. Bio-Med. Eng.* 22: 361-369.
- Merrin E., Meek L.P., Floyd T.C., Callaway E. (1990) Topographic segmentation of waking EEG in medication-free schizophrenic patients. *Int. J. Psychophysiol.* 9: 231-236.
- Meyer-Lindenberg A. (1996) The evolution of complexity in human brain development: an EEG study. *Electroencephalogr. Clin. Neurophysiol.* 99: 405-411.
- Michael D., Houchin J. (1979) Automatic EEG analysis: a segmentation procedure based on the autocorrelation function. *Electroencephalogr. Clin. Neurophysiol.* 46: 232-235.
- Moussavi Z.M.K., Rangaraj R.M., Bell G.D., Frank C.B., Ladly K.O., Zhang Y.T. (1996) Screening of vibroarthrographic signals via adaptive segmentation and linear prediction modelling. *IEEE T. Bio-Med. Eng.* 43: 15-23.
- Newhouse S., Ruelle D., Takens F. (1978) Occurrence of strange attractors near quasiperiodic flow on  $T^m$ ,  $m \leq 3$ . *Commun. Math. Phys.* 64: 35-40.
- Nuwer M.R. (1988) Quantitative EEG: I. Techniques and problems of frequency analysis and topographic mapping. *J. Clin. Neurophysiol.* 5: 1-43.
- Osborne A.R., Provencale A. (1989) Finite correlation dimension for stochastic systems with power-law spectra. *Physica D* 35: 357-381.
- Ott E., Grebogi C., Yorke J.A. (1990) Controlling chaotic dynamical systems. In: *Chaos/Xao - Soviet-American perspectives on nonlinear science* (Ed. K. Campbell). American Institute of Physics p. 153-171.
- Palus M. (1996) Nonlinearity in normal human EEG: cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biol. Cybern.* 75: 389-396.
- Parker T.S., Chua C.O. (1989) *Practical numerical algorithms for chaotic systems*. Springer, New York.
- Pascual-Marqui R.D., Michel C.M., Lehmann D. (1995) Segmentation of brain electrical activity into microstates: model estimation and validation. *IEEE T. Bio-Med. Eng.* 42: 658-665.
- Penczek P., Grochulski W., Grzyb J., Kowalczyk M. (1987) The use of multi-channel Kalman filter algorithm in structural analysis of the epileptic EEG. *Int. J. Bio-Med Comput.* 20: 135-151.
- Pijn J.P., Velis D.N., van der Heyden M.J., DeGoede J., van Veelen C.W., Lopes da Silva F.H. (1997) Nonlinear dynamics of epileptic seizures on basis of intracranial EEG recordings. *Brain. Topogr.* 9: 249-270.
- Pradhan N., Sadasivan P.K. (1996) The nature of dominant Lyapunov exponent and attractor dimension curves of EEG in sleep. *Comput. Biol. Med.* 26: 419-428.
- Pregenzer M., Pfurtscheller G., Flotzinger D. (1994) Automated feature selection with a distinction sensitive learning vector quantizer. *Neurocomputing* 11: 19-29.
- Preissl H., Lutzenberger W., Pulvermüller F., Birbaumer N. (1997) Fractal dimensions of short EEG time series in humans. *Neurosci. Lett.* 225: 77-80.
- Pritchard W.S., Duke D.W. (1997) Segregation of the thalamic alpha rhythms from cortical alpha activity using the Savit-Green S-statistic and estimated correlation dimension. *Int. J. Psychophysiology* 26: 263-271.
- Pritchard W.S., Duke D.W. (1992) Dimensional analysis of no-task human EEG using the Grassberger-Procaccia method. *Psychophysiol.* 29: 182-192.
- Pritchard W.S., Duke D.W., Coburn K.L., Moore N.C., Tucker K.A., Jann M.W., Hostetler R.M. (1994) EEG-based, neural-net predictive classification of Alzheimers disease versus control subjects is augmented by non-linear EEG measures. *Electroencephalogr. Clin. Neurophysiol.* 91: 118-130.
- Pritchard W.S., Duke D.W., Krieble K.K. (1995) Dimensional analysis of resting human EEG. II. Surrogate-data testing indicates nonlinearity, but not low-dimensional chaos. *Psychophysiology* 32: 486-491.
- Rapp P.E., Albano A.M., Zimmerman I.D., Jiménez-Montao M.A. (1994) Phase-randomized surrogates can produce spurious identification of non-random structure. *Phys. Lett. A* 192: 27-33.
- Rechtschaffen A., Kales A. (1968) *A manual of standardized terminology, technics and scoring system for sleep stages of human subjects*. Public Health Service, NIH Publication No.204, W Washington, D.C., US Government Printing Office.
- Rényi A. (1971) *Probability theory*. North Holland, Amsterdam.
- Reschenhofer E., Vollmer R., Deistler M. (1987) Segmentation of the EEG by minimization of the total gain in information. *Neuropsychobiology* 18: 155-159.
- Rombouts S.A.R.B., Keunen R.W.M., Stam C.J. (1995) Investigation on nonlinear structure in multichannel EEG. *Phys. Lett. A* 202: 352-358.
- Röschke J., Aldenhoff J.B. (1991). The dimensionality of human's Electroencephalogram during sleep. *Biol. Cybern.* 64: 307-313.
- Röschke J., Aldenhoff J.B. (1993) Estimation of the dimensionality of sleep-EEG data in schizophrenics. *Eur. Arch. Psy. Clin. N.* 242: 191-196.
- Röschke J., Fell J., Beckmann P. (1995a) Nonlinear analysis of sleep EEG in depression: calculation of the largest Lyapunov exponent. *Eur. Arch. Psy. Clin. N.* 245: 27-35.

- Röschke J., Fell J., Beckmann P. (1995b) Nonlinear analysis of sleep EEG in schizophrenia: calculation of the principal Lyapunov exponent. *Psychiat. Res.* 56:257-269.
- Sammer G. (1996) Working-memory load and dimensional complexity of the EEG. *Int. J. Psychophysiol.* 24: 173-182.
- Sanderson A.C., Segen J., Richey E. (1980) Hierarchical modeling of EEG signals. *IEEE T. Pattern Anal.* 2: 405-415.
- Schiff S.J., Jerger K., Duong D.H., Spano M.L., Ditto N.L. (1994) Controlling chaos in the brain. *Nature* 370: 615-620.
- Sharp F., Smith G., Survillo W. (1975) Period analysis of the electroencephalogram with recordings of interval histograms of EEG half-wave duration. *Psychophysiology* 12: 471-479.
- Shishkin S.L., Brodsky B.E., Darkhovsky B.S., Kaplan A.Y. (1997). EEG as nonstationary signal: the approach to analysis on the base of nonparametric statistics (in Russian). *J. Human Physiol.* 23: 124-126.
- Skrylev K.M. (1984) A method of analysis of abrupt changes in the EEG rhythm (in Russian). *Human Physiol.* 10: 333-336.
- Stam C.J., Jelles B., Achtereekte H.A., van-Birgelen J.H., Slaets J.P. (1996) Diagnostic usefulness of linear and non-linear quantitative EEG analysis in Alzheimers disease. *Clin. Electroencephalogr.* 27: 69-77.
- Stam C.J., van Woerkom T.C., Pritchard W.S. (1996) Use of non-linear EEG measures to characterize EEG changes during mental activity. *Electroencephalogr. Clin. Neurophysiol.* 99: 214-224.
- Takens F. (1981) Detecting strange attractors in turbulence. *Lecture notes in mathematics* (898). Springer, Berlin.
- Theiler J. (1986). Spurious dimensions from correlation algorithms applied to limited time-series data. *Phys. Rev. A* 34: 2427.
- Theiler J., Eubank S., Longtin A., Galdrikian B., Farmer J.D. (1992) Testing for nonlinearity in time series: the method of surrogate data. *Physica D* 58: 77-94.
- Wolf A., Bessoir T. (1991). Diagnosing chaos in the space circle. *Physica D* 50: 239-258.
- Wolf A., Swift J.B., Swinney H.L., Vastano J.A. (1985) Determining Lyapunov exponents from a time series. *Physica D* 16: 285-317.