Reviews of Geophysics
In print, November 2005

Systems Theory for Geospace Plasma Dynamics

Dimitris Vassiliadis[1]

[1] ST at NASA/Goddard Space Flight Center

Abstract

This is a tutorial review on systems theory and its applications to space plasma physics and, more broadly, in geophysics. With its basis on the state representation of a plasma, the theory is widely applicable, but is of particular interest for dynamical, nonlinear, or out-of-equilibrium regimes which cannot be represented by traditional microscopic modeling. Two distinct, but closely related branches of the theory are applied when the plasma dynamics is traced to first principles and when it needs to be derived from experimental data. A framework of modeling methods is presented in order of increasing complexity: enumeration of the effective degrees of freedom, measurement of the linear dynamics and stability, and generalization to their nonlinear counterparts. The relation between symmetries in the plasma system and modes in its structure and response is discussed. Signal processing methods are presented, illustrated by examples, and their relative merits and limitations are discussed. The dynamical framework provides a new approach alongside the traditional perturbative and statistical-mechanical methodologies, and is directly relevant to the development of space weather applications.

Contents

1. Introduction[2]

In modeling realistic plasmas, first-principles approaches are not always applicable. Often, the dynamics is due to strongly nonlinear interactions; equilibria do not exist or are unstable; the probability distribution may be nontrivial (e.g., highly non-Gaussian or non-power-law, such as multi-modal) and therefore its moments difficult to calculate, etc. In many of these instances, systems theory provides practical answers in representing the plasma dynamics.

The defining feature of the systems theoretical approach is that the plasma is represented by a time-dependent state vector, comprised of a small number of relevant variables. In principle, the state may

[2] The paper is based on lectures given at the School of Analysis Techniques for Space Plasma Data held in La Londe-Les Maures, France, on October 8-13, 2001.

include all microscopic variables such as particle positions and velocities. Most often, however, the state is not known in detail, and the variables used instead are several moments of the particle distribution and components of large-scale fields. In fact, state variables are fairly general functions of experimental measurements as long as they are demonstrably useful in representing the system dynamics. Construction of the state space will be discussed in detail in Sections 2.3 and 4.2.

The number of state variables is small compared to the number of degrees of freedom of the actual system: it turns out that for many geospace plasmas, the large-scale dynamics may be represented to satisfactory accuracy by a few tens of variables or less. The physical justification for the truncated description is given in Section 2. The emphasis is on modeling the physically, or at least phenomenologically, significant variables which can account for the observed variability (Sections 2.1 and 5.2-5.3).

The state space approach can be related to other methodologies. If the dynamics is known from first principles, as the number of state variables increases, the approach yields a mesoscopic or microscopic model. Under these conditions, a state space model with many degrees of freedom tends to a full plasma simulation. On the other hand, if we are interested in the average properties rather than detailed time-dependence, we can take appropriate phase or ensemble averages of state space model results.

A key characteristic of many system theoretical methods is their geometrical interpretation. As the state vector evolves in time, it traces out a trajectory in a vector space –the state space. Dynamical characteristics such as rates of growth or decay, inductive timescales as part of a response to external shocks, etc., produce spatial patterns in the state space. Properties of the state space are closely analogous to those of the phase space for Hamiltonian dynamics [Goldstein].

Having these insights as the background, we now resume the discussion why a systems theoretical approach is useful for representing space plasma environments, as it has been useful in other fields (engineering, econometrics, biology, etc):

Equilibria and dynamics of geospace plasmas. Geospace plasmas evolve towards nonlinear equilibria, or even towards far-from-equilibrium regimes where traditional perturbation (quasilinear) theory is of little use. Furthermore, several aspects of the dynamical repertoire observed in geospace are reproduced succinctly by relatively "simple", nonlinear, time-dependent models.

Open plasma systems. Dynamics and stability studies are often based on assumptions of relatively simple geometries and boundary conditions, time-independent external forcing, and a small number of dominant interactions. These assumptions are useful for understanding the basic physics of the system. Realistic plasmas, however, interact continually with their environments (plasmas, planetary atmospheres, or lithospheres) through time-dependent inputs and at complex spatial boundaries. As a result, the observed behavior cannot be modeled with simple assumptions. Instead of using a fully kinetic representation of the interaction between plasma systems, one can use a data-based input-output model (Sections 4.3, 5.2, and 5.4).

Complex temporal and parametric behavior. Since its evolution from classical mechanics, systems theory has encompassed a large number of concepts and tools which have proven useful in understanding complex dynamics. The concepts include dynamical chaos, bifurcation theory, catastrophe theory, multifractals, self-organization and self-organized criticality, etc. These dynamical elements have been identified in fluid and plasma systems in the laboratory . It is highly likely that they can be identified in the dynamics of geospace plasmas as a growing number of recent studies indicates (Section 5).

Furthermore, historically discoveries in space plasma physics and astrophysics follow increases in observational or analytical capabilities. With extensions into inverse modeling and data analysis, systems-theory provides an array of methods for visualizing, quantifying, and representing plasma measurements.

Space weather applications. The recent availability of real-time monitoring of geospace environments, and especially of the interplanetary medium, has transformed static, physical models to dynamic, predictive ones. This technological development enabled the creation of national space weather programs [Behnke et al., 1995; 1997], repeating the historic progress in meteorology and oceanography. In those fields, the combination of systems theory, large-scale fluid simulation and intensive, real-time data acquisition led to data assimilation in the 1970s and 1980s, a development which is being repeated in space weather since the mid-1990s.

Historical highlights.

The broad scope of today's systems theory is in contrast with its esoteric origins at the end of the 19th century when Poincaré studied motion and stability in the gravitational potential of planets and few-satellite systems [Poincaré, 1957]. Interestingly, those first studies focused on the space environment from the perspective of celestial dynamics. One of Poincaré's major discoveries was that for certain simple, general potentials, no integrable (closed-form) solution could be found, which means that motion is aperiodic, analytically intractable, and therefore unpredictable in the long term, while still being bounded. For some regimes, motion is thoroughly perturbed that a particle visits every neighborhood of the phase space to it ergodically, i.e. with approximately equal probability. The aperiodic, unpredictable behavior of simple deterministic systems was a paradigm shift from the Newtonian clockwork-like model of the universe [Stengers and Prigogine, 1984]. Poincaré's discoveries were not extended further until decades later, when electronic computers facilitated the calculation of complex state space trajectories.

In the mid-20th century, the work of Mandelbrot focused attention on the significance of fractality (self-similarity) which occurs in natural and artificial systems as diverse as mountains, coastlines, and financial markets [Mandelbrot, 1983]. Lorenz showed how a truncated, few-mode, description of a standard atmospheric circulation model can become intrinsically unstable in the absence of external time-dependent perturbations [Lorenz, 1963]. In contrast to Poincaré's gravitational systems which preserved a Hamiltonian (energy) function, this was a dissipative system where energy had to be continuously replenished and information was rapidly lost. Even though the system was deterministic rather than stochastic, the nonlinearity led to deterministic chaos, that is, lack of predictability after a short decorrelation time. For a while chaos became a likely candidate to explain fluid (and plasma) turbulence. Routes to chaos, ostensibly analogous to transitions from laminar to turbulent motion, were discovered and, in the process, sophisticated analysis tools were developed [Feigenbaum, 1978; Ruelle and Takens, 1971]. Nonlinear relaxation oscillators were discovered in a variety of systems including thermionic discharges and plasma devices. Difference equations of few degrees of freedom exhibited complex phase-space patterns, broadband power spectra, and intertwined stable and unstable equilibria [Li and Yorke, 1981].

In the subsequent decades, nonlinear systems theory became an underpinning of modern engineering analysis including system identification, stability, and eventually control. In one particular application, space engineering returned to its Poincaré origins: in interplanetary mission design, trajectories need to be complex enough to enable long-term observations, yet minimize the amount of energy expended. Advanced system-control methods in combination with high-resolution gravitational maps of the solar system enabled the design of a network of highly complex trajectories dubbed the "gravitational superhighway" [Howell et al., 1997; Lo et al., 2001].

In space physics, the first applications of systems theory were made in the early 1980s. Modeling the particle transport and stability in a curved magnetic field approximating the geometry along the Earth's magnetotail axis [Chen and Palmadesso, 1986; Büchner and Zelenyi, 1989; Burkhart et al., 1992]. Depending on the field configuration and the initial conditions, there were very different properties in transport and stability. At about the same time, system identification methods were used to calculate FIR ("linear prediction") filters from geomagnetic index data such as D_{st} [Iyemori et al., 1979; Bargatze et al., 1985]. Once the filters were calculated, they could be used to predict a different dataset of the index. Since those early applications, systems theory has been used in many different ways in the four plasma approximations (particle, kinetic, fluid, and global-planetary). Therefore the "system" that we will refer to below can represent the dynamics of a test particle, the moments of a plasma distribution, a number of magnetic-field modes, the large-scale dynamics of the magnetosphere, or any other from a broad selection of plasma dynamics.

2. Fundamental concepts

While the applicability and domain of validity of systems theory in plasma physics may be evident in some cases, it may not be in others. Therefore this section discusses several basic notions in order to introduce and provide a physical foundation for the theory. The first two parts discuss the relevance of the approach, and the third one introduces the construction of a state space from experimental data.

2.1 Plasmas organize in modes

In modeling the dynamics of a plasma, a useful starting point is the number of its degrees of freedom (df). For a structureless system with n particles, which do not possess internal degrees of freedom, $df = 6N$. Even for low-density systems such as geospace plasmas, this is a staggering number. The plasma sheet in the Earth's magnetotail, for instance, extends over a volume of ~70x20x6 R_E^3 with an average ion density of 0.1 particles/cm^3, or a total of ~10^{30} protons.

A space plasma, however, is observed to behave as if having different numbers of degrees of freedom at different times. This is especially true at times of high activity, when it turns out that a much lower number than 6N is sufficient to reproduce or predict the large-scale features. This lower number will be called the number of effective degrees of freedom, df_{eff}.

In order to attempt to estimate the df_{eff} of the plasma sheet, one needs either multi-point in situ measurements or global, remote-sensing measurements. The former is not available until the advent of multi-spacecraft constellations. However, the dynamics, such as plasmoid ejection and flow acceleration, can be indirectly monitored by examining the UV emission of the nightside auroral oval and polar cap measured by Viking, Polar and other spacecraft [Ieda et al., 2001]. The two-dimensional height-integrated emission pattern is generated by plasma sheet electrons precipitating into the ionosphere along magnetic field lines. During quiet times, the emission pattern is characterized by a large number of low-amplitude, short-lived flashes, which number several thousand over the entire high-latitude region viewed at the Polar/UVI resolution (70 km). During storms and substorms, however, the precipitation is organized in large-scale arcs and the luminosity variations of the auroral oval and polar cap can be represented by df_{eff}~10-100 variables, depending on the event intensity. The df_{eff} reduction in the luminosity pattern is representative of a reorganization of the degrees of freedom in the plasma sheet proper.

In a complex system such as the plasma sheet (or a fluid, lattice, or biological system), intensive variables such as energy density, momentum, temperature, etc., are not distributed uniformly. From a different perspective, it is information (or entropy) that is nonuniformly distributed [Tsallis, 1988]. In either description, the relevant variables are heavily weighted in some degrees of freedom and are negligible for others. Therefore the number of effective degrees of freedom, df_{eff}, is much lower than the df estimated at the microscopic (particle) level.

Reduction of degrees of freedom.

The reduction in df_{eff} at several length and time scales is the direct result of collective effects in the plasma system. The first significant reduction occurs at scales much less than a particle's mean free path. Thermal motion limits the effects of individual charged particles at distances larger than the Debye length λ_D (Fig. 1). For a plasma in a volume of linear scale L at uniform density n_0, the effective number for a macroscopic description ($r \gg \lambda_D$) is bounded by

$$df_{eff} < \frac{6N}{N_D} \sim \frac{N}{n\lambda_D^3} \sim \left(\frac{L}{\lambda_D}\right)^3$$

Second, in the presence of a magnetic field, adiabatic invariants limit a particle's motion [Baumjohann and Treumann, 1996] and lead to a second reduction in df_{eff}. At the MHD approximation, the field constraints correspond to the plasma being frozen into and co-moving with the field. In regions of finite resistivity the frozen-in condition does not hold and therefore df_{eff} increases locally.

A third, probably more interesting, reduction occurs dynamically when instabilities are excited (Sec. 4.1). Instabilities organize a part of the particle distribution into a spatial structure, or mode, of a certain duration (the inverse of the instability decay rate).

Modes and symmetries.

Thus the effective degrees of freedom are associated with modes, which are generalized nonlinear spatiotemporal equilibria. Time-stationary modes include Grad-Shafranov solutions, their extension in the low-beta case to force-free equilibria such as plasmoids and flux ropes, and in the high-beta case hydrodynamic structures, such as eddies. Time-dependent modes are discontinuities and waves, which also lead to a reduction in the effective degrees of freedom. In real-world plasmas, modes are complex disturbances in space, time, and activity level. For example, magnetospheric substorms, convection bays, and pseudobreakups are large-scale nonlinear modes of the magnetotail current, respectively (e.g., [Horton and Doxas, 1996]; see Sec. 5). During magnetic storms, there is a coherent response to passing interplanetary structures in both the ring current [Valdivia et al., 1999] and the radiation-belt relativistic electron flux (e.g., [Baker et al., 1999]. Interplanetary inputs excite different modes in the magnetotail [Bargatze et al., 1985] and the radiation belts [Vassiliadis et al., 2002; 2003].

The degrees of freedom that participate in the same mode have similar dynamics except for a phase differences. In order to model the coherent motion of the particles involved in the formation of a mode, one can introduce a symmetry in the equations of motion. Thus different modes correspond to distinct local symmetries in the plasma distribution.

For example, in a plasma of N particles, consider a group of N_g particles that forms a mode (Fig. 2). Since the N_g particles are subject to the same coherent dynamics, the only distinction among them is a phase difference. If we decide to neglect the phase variable in the modeling, particlaes in the group become indistinguishable, or symmetric with respect to exchange. The effective degrees of freedom are reduced to $df_{eff}=6(N-N_g)+N_m$, where N_m is the number of degrees of freedom needed to represent the mode. Typically $N_m<<6N_g$, and therefore there is a reduction in df_{eff}.

For practical modeling considerations then, one can see that if the modeling is based on a few experimentally measured variables, the similarity between the effective degrees of freedom provides significant flexibility in choosing which variables, or functions thereof, are used in modeling. The remaining degrees of freedom are lower in amplitude and are ignored.

Thus a mode-based description, when available and valid, is of great advantage in modeling and understanding plasma activity.

Mode formation.

One can see that there are three basic scenarios for forming modes:

A. An external driver, such as an electromagnetic field, imposes structure and dynamics on the plasma. In the case of strong driving, the time dependence of the plasma, in terms of an observed variable or output O(t), is similar to that of the driver, or input, I(t) up to a phase difference.

$$O(t) \sim I(t) e^{i\Delta\phi}$$

A familiar example is the condition of the high-latitude ionospheric potential, which is determined by the large-scale magnetospheric electric field and electron precipitation. Other large-scale effects, such as ionospheric feedback and plasma-neutral interactions with the upper atmosphere, although important, can be neglected at first approximation.

B. In certain cases, plasmas display spontaneous self-organization into modes in the absence of time-dependent driving. Self-organization appears in MHD and hydrodynamic turbulence [Hasegawa, 1985]. As the system's internal energy is dissipated, other global variables decrease more slowly or even grow with time. Such variables are called rugged invariants [Matthaeus and Goldstein, 1982] and their modes are essential elements of the system description. Self-organization can produce dynamical instabilities, nonlinear waves, and chaos [Hasegawa, 1985; Ortolani, 1989].

Example: 3-D MHD turbulence. The unstructured solar wind (i.e., in the absence of transient structures such as shocks, coronal mass ejections, and streams) is understood reasonably well in terms of near-ideal MHD turbulence. Three-dimensional spectral simulations of homogeneous, compressible turbulence have shown that there exists a rugged invariant, namely the helicity

$$H = \int_V (\mathbf{A} \cdot \mathbf{B})\, dV$$

where **A** is the vector potential, $\nabla \times \mathbf{A} = \mathbf{B}$ [Ting et al., 1986; Ghosh and Matthaeus, 1990]. The evolution of the helicity is slower than that of the total energy. The energy is transferred from smaller wavenumbers to larger ones, or large eddies in the magnetofluid break up into smaller ones, and the process continues until the wavenumber reaches the dissipation range. This type of transfer of a conserved quantity in wavenumber space is called a direct cascade, and is conceptually similar to energy cascade and dissipation in neutral fluid turbulence [Kraichnan, 1967].

Simultaneously, however, helicity may grow. Helicity is transferred from higher to lower wavenumbers in a so-called inverse cascade [Matthaeus and Montgomery, 1984]. Therefore, the plasma self-organizes in the sense that many degrees of freedom spontaneously form large-scale helicity eddies. The eventual outcome of helicity growth or decay depends on the initial conditions (Fig. 3). More generally, selective decay and self-organization in modes is evident in the variables associated with the rugged invariant, but not in those associated with the energy. The dichotomy has been eloquently expressed in Hasegawa's review [1985].

C. In actuality the evolution of the system is determined by both the driver (if present) as well as the internal processes.

$$\frac{dO}{dt} = f\left[O(t), I(t)\right] \tag{2.1}$$

An example is solar wind-magnetosphere coupling where the magnetospheric storm or substorm response is represented as a nonlinear oscillation or instability excited by the interplanetary driver [e.g., Klimas et al., 1994]. Sections 4 and 5 discuss ways of modeling effects of the type (2.1). The storage and dissipation of the free energy can be described in terms of critical phenomena theories including phase transitions [Sitnov et al., 2000], the renormalization group [Chang, 1992], and self-organized criticality [Chapman et al., 1998].

2.2 Plasmas are fractal

Work by Mandelbrot [1967] and others has shown that the spatial structure of natural systems has characteristic power-law scalings: structure is organized in self-similar, or scale-invariant, patterns called fractals. Geophysical systems share these properties [Turcotte, 1997] and so do geospace plasmas.Power-law distributions are well-known from many space plasma contexts, for instance in turbulence or energetic-particle spectra. In the last two decades, however, power laws have been found in the spatial and temporal distributions of a wide variety of space plasmas such as solar flares, geomagnetic field fluctuations, and ionospheric UV emissions. In these systems the distribution is self-similar in the sense that an increase in resolution, spatial or temporal, depending on the system, produces a distribution statistically similar to the original one [Mandelbrot, 1983]. For many of these systems, the structure of the state space is self-similar as well. The scale invariance can be traced to symmetries in the system equations, and is akin to the older concept of mechanical similarity . The fractal paradigm has been extremely useful in geophysics [Turcotte, 1997] with perhaps the archetypical self-similar scaling being the Gutenberg-Richter relation for earthquake occurrence.

Example: The lightning discharge. A lightning bolt is a plasma with a fractal charge distribution. The dendritic (tree-like) path of the cloud-to-ground stroke is composed of a few major branches connected to increasingly smaller arcs (Figure 4). The branches and arcs are created as fluctuations in the electric field are produced by the dielectric breakdown and feed back to it until the final path is formed. Dendritic and tortuous patterns are common in dielectric discharges and can be explained by a nonlinear feedback of the field [Niemeyer et al., 1984].

Self-similarity in lightning is quantified through image analysis [Kudo, 1998] and modeling [Vecchi et al., 1994]. One method is to measure the extent to which a two-dimensional pattern f(x,y) at resolution δr is statistically similar to its magnification by a factor A (Fig. 5). At scale $r_i=A^i\delta r$ the number of bright pixels above an intensity threshold in a given wavelength is $N_i = N_i\left(r_i\right)$ The lightning pattern is self-similar if there is a power-law scaling between the pixel number and the resolution:

$$N_i\left(r_i\right) = c r_i^{d_0} \tag{2.2}$$

where c is a constant.

The capacity dimension d_0 indicates the degree of self-similarity and is one of many scalings of fractal distributions. This dimension indicates how well the structure fills out the surrounding three-dimensional space. It is estimated from measurements at two length scales r_i and r_{i+1} as

$$d_0 = \frac{\ln\left(N_{i+1}/N_i\right)}{\ln\left(r_{i+1}/r_i\right)} \tag{2.3}$$

Typical values of d_0 for cloud-to-ground lightning measured from photographic images are 1.3-1.4, meaning that the distribution of the discharge has a dimension between a Euclidean line and a two-dimensional plane (Fig. 6). In practice, estimates of (2.3) from a broad range of scales r_i are averaged or compared. The ratio between the smallest and largest scales r_i should ideally be >3 orders of magnitude although in practice it is limited by the image resolution and the cloud altitude.

Thus a fractal dimension represents a statistical property of the electric field in a succinct way unmatched by traditional techniques, such as Fourier analysis. The dimension is then related to properties of the atmospheric medium during the storm and the ground environment, that is, the boundary conditions.

In addition to the spatial distribution, other properties of a fractal plasma are also self-similar: A time-dependent branched discharge acts as an antenna with a radiated power and temporal signature characterized by the power laws of the branching [Puente-Baliarda et al., 1998]. Intra-cloud discharges emit electromagnetic pulses whose far-field emission pattern is highly inhomogeneous [Valdivia et al., 1998]. Its absorption at ionospheric altitudes accelerates electrons and produces secondary, red-sprite lightning, which also has fractal structure [Valdivia et al., 1997;]. The spatial dimension of a discharge is directly related to the "dimension" of the time series of the radiated field [Vecchi et al., 1994]. Conversely, measuring properties of a field or a time series provides information about the distribution of the original charge or current sources.

Fractal distributions of charges and currents are ubiquitous in geospace. Prominent examples include the distribution of electron precipitation in the auroral ionosphere as revealed by their UV emission [Lui et al., 2000; Uritsky et al., 2002] and the structure of current sheets in the Earth's magnetotail [Milovanov et al., 2001]. The interplanetary medium, the long-term properties of the solar wind reveal that it is organized as a fractal structure [Burlaga, 2001].

Fractals in time.

Fractal behavior is not limited to spatial patterns. The activity of plasmas varies in self-similar sequences as they interact with their environment, undergo internal dynamics, or radiate. In addition, traversal of a spatially self-similar plasma system produces fractal time series as measurements. For a given system the degree of self-similarity in time is in general different than the degree of self-similarity in space (e.g., Vecchi et al. [1994]). The difference between the two depends on the dispersion relation $\omega=\omega(k)$.

The best-known temporal fractals due to geospace processes are time series of geomagnetic indices.[3] Other important fractals include the time distribution of auroral luminosity patterns due to electron precipitation in UV [Lui et al., 2000; Uritsky et al., 2002] and the distribution of solar flares in the hard X-ray spectrum [Lin et al., 1984; Vlahos et al., 1995]. In addition the distribution of current disruption

[3] A geomagnetic index is a measure of the ground magnetic effects due to one or more current systems [e.g., Kivelson and Russell, 1995]. Indices provide long-term, continuous measurements of activity and are used to characterize events of interest thereby facilitating comparisons and statistical studies. Their advantages and limitations have to be assessed critically for each application [e.g., Baumjohann, 1986].

signatures in the near-Earth magnetosphere appears to be fractal [Lui et al., 1988; Ohtani et al., 1995, 1998] although analyses of in situ measurements contain an element of space-time ambiguity due to spacecraft motion.

Spatial patterns are often self-affine: the value of the fractal dimension measured in one direction is different from the value as measured in a perpendicular direction. Temporal patterns have this property as well: amplitude variations are correlated with, but not statistically identical to variations in time development. Consider the graph of the auroral electrojet index[4] AE(t), at a resolution ΔAE and with measurements taken at resolution Δt. Index fluctuations are denoted by ΔAE(t; Δt). A magnification by r in both the time and amplitude resolutions provides new information in the form of a new set of fluctuations, ΔAE(t; rΔt). The graph is self-affine if, after a magnification by a factor r, the original fluctuations are related to the new as

$$\left|\delta AE\left(r\Delta t\right)\right| = r^{H_a}\left|\delta AE\left(\Delta t\right)\right| \tag{2.4}$$

where H_a is the Hausdorff dimension. For $H_a \neq 1$ fluctuations in the index grow at a different rate than the changes in time. In a similar manner, self-affine spatial fractals are marked by an anisotropic growth of fluctuations with scale.

Self-affinity is measured by the variance of the time series. We can estimate a sample variance assuming that the sample collected over a time T is representative of the population. The variance, or squared standard deviation, of an index such as AE is defined as

$$\sigma_{AE}^2 \equiv \int_{AE}\left(AE(t)-\overline{AE}\right)^2 dAE \triangleq \frac{1}{T}\int_0^T\left(AE(t)-\overline{AE}\right)^2 dt \tag{2.5}$$

where $\overline{AE}$ denotes the time average over the interval T. The variance of a self-affine time series scales with time as:

$$\sigma_{AE}^2(T) \sim T^{2H_a-1}$$

2.3 The state representation of a plasma system

A state space is the basic representation of a physical system in the systems approach [Abarbanel et al., 1993]. The state is constructed either from the physical equations, when available, or from the observed variables. For analysis, it provides significant advantages compared to spectral and covariance-based methods. In constructing a state space from time series data, a procedure also called embedding, temporal characteristics (amplitude, phase, discontinuities, stability, etc.) produce spatial counterparts (size, radial/angular position, region boundaries, etc.). For low-dimensional state spaces or their sections (subspaces), visualization of the spatial features can be a valuable aid to the modeler.

The state **x** of a plasma system is a vector which represents its activity level and is comprised of the system variables.

$$\mathbf{x}(t) = \left[x_1(t), x_2(t), \ldots, x_{m_0}(t)\right] \tag{2.6}$$

The number of system variables is m_0. The time-dependent state **x**(t) is observed via a measurement process H which yields an output **y**(t) of lower dimension. For example, the following system:

$$\begin{aligned}\frac{d\mathbf{x}}{dt} &= \mathbf{A}\cdot\mathbf{x} \\ y(t) &= \mathbf{H}\cdot\mathbf{x}(t)\end{aligned} \tag{2.7}$$

[4] The auroral electrojet index AE provides a way to estimate the activity of the intense horizontal ionospheric currents, or electrojets, in the auroral zone (65-70°N). The currents intensify either due to convection or to substorm development following transfer of momentum and energy from the solar wind to the magnetosphere-ionosphere system.

has linear time-independent dynamics, **A**, and a scalar output, y(t).

States **x**(t) at successive time instances trace a trajectory in the state space. The correspondence between temporal dynamics and state-space structure is a key element of systems theory and the basis for two complementary approaches, forward and inverse modeling. In the former, the evolution equation of the system is known and explicitly used to understand its properties. In the latter, the source of information about the system is a series of experimental measurements, represented by y(t) above, and the dynamics and other properties must be extracted from that information.

Forward modeling.

Integrating an evolution equation like (2.7) is used to form a state space. Geometrical methods are used to quantify direction, length, and curvature of the trajectory, which are then related to physical parameters such as instability growth rates and energy dissipation rates. These methods are more powerful in pattern identification and modeling of irregular, "noisy" time series produced by nonlinear systems than conventional linear (spectral) time series analysis.

State spaces can be constructed for a variety of different plasma systems. The familiar phase space of classical mechanics [Goldstein, 1980] is a type of state space. For a test particle moving in a time-independent one-dimensional potential V(x), for instance, the phase space is (x,p) where p is the particle momentum (Fig. 7). More generally, the state of an N-degree of freedom system is a 2N dimensional vector

$$\mathbf{x}(t) = \left[q_i(t), \ldots, p_i(t)\right],\ i = 1, 2, \ldots, N$$

comprised of action variables $q_i(t)$ and conjugate momenta $p_i = \dfrac{\partial L}{\partial \dot{q}_i}$, where L is the Lagrangian of the system. For Hamiltonian systems there is a quantity, the Hamiltonian which is given by $H = \sum_{i=1}^{N} p_i \dot{q}_i - L$ and represents the total (potential and kinetic) energy. The state evolves according to the equations of motion:

$$\dot{q}_i = \frac{\partial H}{\partial p_i}$$

$$\dot{p}_i = -\frac{\partial H}{\partial q_i}$$

A Hamiltonian can be written for charged-particle motion in a curved static field such as a Harris sheet. Since the Hamiltonian is conserved, its initial value determines the geometry and stability of the motion. Fig. 8 shows a special type of cross-section of the phase space, called a Poincaré section [Chen and Palmadesso, 1986; Büchner and Zelenyi, 1989]. The footpoints of quasiperiodic trajectories produce regular structures, seen as closed curves, or islands, on the Poincaré section while irregular trajectories form the surrounding shaded region. Hamiltonian systems are conservative which means the volume of a phase space element is constant over time according to Liouville's theorem [Goldstein, 1980].

State space structures are very different for dissipative systems, where energy has to be provided externally. A model of the large-scale magnetospheric activity is based on the electrodynamics of the magnetotail [Klimas et al., 1992]. Magnetic reconnection on the dayside produces a excess of open magnetic flux which is stored in the magnetospheric lobes and causes the tail current to intensify. Depending on the rate of the input, the stored energy either decays slowly as the excess flux recombines and is convected back to the dayside, or is released explosively during a dipolarization of the tail field. If the dipolarization occurs across the entire width of the tail, the event is called the magnetospheric substorm. The model has three dynamic variables: the lobe flux $\Phi = \int_A \mathbf{B} d\mathbf{A}$ where **B** is the lobe field and A is the lobe cross-section; the cross-tail electric field E_{tail} whose difference with the interplanetary electric field determines the time rate of change of the lobe flux,

$$\frac{\partial \Phi}{\partial t} = -\left(E_{tail} - E_{SW}\right) \tag{2.8}$$

and the cross-tail current j_{tail}, related to the E_{tail} through the ion momentum equation, and to the flux through the lobe field it produces. Thus a particular choice for the state space is $\mathbf{x}(t) = \left[\Phi_{lobe}(t), E_{tail}(t), j_{tail}(t)\right]$. Fig. 9 shows the projection of the state space on the (E_{tail}, dE_{tail}/dt) plane where the model's substorm phases are indicated. The [Klimas et al., 1992, 1994] model and its descendants will be discussed in Section 5.2.

Inverse modeling and embedding methods.

For many systems of interest, it is possible to reconstruct the state space from the observed y(t) and calculate evolution equations such as (2.7). The applicability of this approach is specified by embedding theorems.

The first step is to produce state variables from the observations y(t). There is significant freedom as to the choice of the variables. If some of the original state variables are not measured, they can be replaced by functions of the observed variables. The equivalence between state spaces produced by different combinations is given by embedding theorems [Sauer et al., 1991] as discussed in Section 4.2. The procedure of generating state variables from observations is called reconstruction of the state space.

The most popular method of state space reconstruction is time delay embedding. Consider that the observation vector $\mathbf{H} = [1, 0, 0, \ldots]$ in Eq. (2.7) and therefore $y(t) = x_1(t)$. The reconstructed state is made up of successive measurements of y(t):

$$\mathbf{x}(t) = \left[x_1(t), x_1(t-\tau), x_1(t-2\tau), \ldots, x_1\left(t-(m-1)\tau\right)\right] \tag{2.9}$$

The vector represents the recent history of the system starting from (m-1) steps in the past; by convention more recent observations are listed earlier than more distant ones. Generally the reconstructed space dimension is $m \geq m_0$.

The representation of a complex plasma system in terms of a relatively low-dimensional state space, whether in the original form, Eq. (2.6), or the reconstructed one, Eq. (2.9), is a bold modeling approach. It relies on the fact that, in many cases (but not always), collective effects lead to mode formation and a reduction in df_{eff}. State space reconstruction methodologies are discussed in Sections 3.4 and 4.2, and prediction-based tests that are used to determine the optimal model for a given application are given in Sec. 4.4.

3. Counting the effective degrees of freedom

3.1 Significance

Estimating df_{eff} is a basic step in the systems theory framework and is useful in determining the type of modeling to be applied. If $df_{eff}<10$, one can use a low-order model as discussed in Sections 4 and 5. For a large number of degrees of freedom (>100), simulation or statistical-mechanical approaches are preferable. In some problems, estimation is the most that can be, or needs to be, accomplished. Also in certain cases the methods used give some guidance for the physical interpretation of the degrees of freedom. A low number of degrees of freedom for a system with broadband spectra can be a signature of chaos, provided that external driving effects are accounted for and the numerical method can be interpreted in a unique fashion. The following example serves to introduce the issue.

Example: Langmuir oscillations.

Consider a one-dimensional plasma featuring cold (monoenergetic) electrons and heavy ions, forming a uniform positive background. The continuity and momentum equations for the electrons are:

$$\frac{\partial n}{\partial t}+\frac{\partial(nv)}{\partial x}=0 \tag{3.1}$$

$$m_e n\frac{\partial v_1}{\partial t}=-enE-\gamma nv \tag{3.2}$$

The last term in Eq. (3.2) represents a decelerating drag of nearby ions and is small compared to the field acceleration.[5] The electric field is given by Poisson's equation:

$$\frac{\partial E}{\partial x}=-\frac{e}{\varepsilon_0}n \tag{3.3}$$

At t=0 an electric field impulse $E_0\delta(t)$ induces Langmuir, or plasma, oscillations (Fig. 10) [Chen, 1984]. The system variables, the electron density and velocity, are perturbed from their their average values:

$$n(x,t)=n_0+n_1(x,t)$$
$$v(x,t)=v_0+v_1(x,t)=v_1(x,t)$$

where n_0 is constant, $|n_1|\ll n_0$, and $v_0=0$. In the linear approximation the system equations are:

$$\begin{aligned}&\frac{\partial n_1}{\partial t}+n_0\frac{\partial v_1}{\partial x}=0\\&m_e\frac{\partial v_1}{\partial t}=-eE-\gamma v_1\\&\frac{\partial E}{\partial x}=-\frac{e}{\varepsilon_0}n_1\end{aligned} \tag{3.4}$$

and can be combined into:

$$\frac{\partial^2 n_1}{\partial t^2}+\frac{\gamma}{m_e}\frac{\partial n_1}{\partial t}+\omega_p^2 n_1=0 \tag{3.5}$$

where $\omega_p^2=\frac{n_0e^2}{\varepsilon_0 m_e}$ is the electron plasma frequency. A Langmuir probe at fixed $x=x_L$ measures a current which is proportional to the density perturbation $n_1(x_L; t)$. Thus the system output varies according to

[5] Part of the electron energy is slowly transferred to the ions to induce ion-acoustic waves (see [Chen, 1984; Sturrock, 1994] for more details). The presence of waves decreases df_{eff} further.

$$\frac{d^2 n_1}{dt^2}+\frac{\gamma}{m_e}\frac{dn_1}{dt}+\omega_p^2 n_1=0 \qquad (3.6)$$

The question is, what is the number of effective degrees of freedom, df_{eff}, for this plasma system. As before, it is addressed in two different contexts, forward and inverse modeling.

Forward modeling.

Analysis of (3.6) with the ansatz $n_1 \propto e^{i\omega t}$ leads to a dispersion relation which is quadratic in the complex frequency ω. The oscillation frequency and growth rate are

$$\omega_r = \pm\sqrt{\omega_p^2-\left(\frac{\gamma}{2m_e}\right)^2}$$

$$\omega_i = -\gamma/2m_e$$

Only time-independent and growing modes contribute to the degrees of freedom because the assumption is that the system has reached equilibrium. Thus there are two effective degrees of freedom, represented by, e.g., (n_1, dn_1/dt), (v_1, dv_1/dt), or other combinations. Other degrees of freedom can be expressed as a function of the effective ones with an additional phase difference.

Inverse modeling.

Methods for estimating df_{eff} from experimental data are discussed next. They are divided into spectral and covariance-based methods. State space methods are a part of the latter group, but are discussed in a separate section.

3.2. Spectral methods

These methods project the experimental time series on a basis of orthogonal vectors, and df_{eff} is estimated as the number of significant components, i.e., those whose magnitude exceeds a certain value. Most often the methodology involves a Fourier or spherical harmonic analysis of the data [for the expansions see: Arfken and Weber, 1995; Press et al., 1992; Glassmeier et al., 1995], but recently methods based on wavelets are increasingly used [De Wit, 2002].

In Fourier analysis, an experimental time series x(t) at steps t=1,2,…T (for example the density measurements $n_1(t)$ of Eq. (3.6)) is written as

$$x(t)=\frac{1}{2\pi}\int_{\omega=-\infty}^{\infty} X(\omega)e^{i\omega t}d\omega$$

where the Fourier component at frequency ω is given by

$$\tilde{X}(\omega)=\sum_{t=1}^{T}x(t)\mathrm{e}^{-i\omega t}$$

and defines a time series $\tilde{X}(\omega)e^{i\omega t}$. The significance of a Fourier component at frequency ω is measured by the power spectral density around a narrow frequency interval

$$P(\omega)=\frac{1}{2}\int_{\omega-\delta\omega}^{\omega+\delta\omega}\left|\tilde{X}(\omega)\right|^2 d\omega \qquad (3.7)$$

The spectral peaks obtained from a representative time series can be used to count and rank the df_{eff}, with each peak representing two degrees of freedom. Any background spectral power is subtracted from (3.7).

Other spectral methods.

Windowed transforms [Press et al., 1992] and wavelets [Chui, 1992] can be used for df_{eff} estimation in nonlinear or nonstationary systems. The maximum window size is determined from the physical or dynamical time scales of the plasma. Wavelets have exact closure relations so they are useful for developing empirical equations for the dynamics such as wave-wave coupling [De Wit, 2002].

Advantages and limitations.

Spectral methods work best for stationary time series with low levels of noise. Spectral analysis relies on the existence of symmetries either in the system geometry, the form of the interaction, or its boundaries. Symmetries lead to a quantization of the allowed solutions and to the orthogonality property. Spectral approximations may be used for weakly nonlinear or nonstationary systems. In those cases, however, spectral methods tend to overestimate the number of modes because additional modes may appear due to mode coupling. High-order spectral methods are used to represent strongly nonlinear coupling between modes or between the system and its driver [Lefeuvre et al., 1995; De Wit et al., 1995, 1999; Balikhin et al., 2001a,b;]. Nonlinear mode coupling typically leads to broadband spectra and the apparent existence of many degrees of freedom.

Applications.

Traditionally, spectral methods have been particularly useful in the study of pulsations and field line resonances in the inner magnetosphere [Anderson et al., 1990; Takahashi et al.] and on the ground [e.g., Anderson et al., 1990; Chi and Russell, 1998; Engebretson et al., 1998]. They are useful in the identification of wave mode, both in situ and remotely, [e.g., Song et al., 1994; Glassmeier et al., 1995]. In those cases the system order is the number of significant modes in the plasma [Motschmann and Glassmeier, 1995].

Analysis of ground-based magnetometer data has incorporated several spectral-matrix techniques [Samson, 1983], notably the cross-phase spectrum [Chi et al.].

3.3. Statistical methods

While the usefulness of spectral methods relies on the existence of known symmetries in the spatial structure of the plasma or in its interactions, in many cases this information is not available clearly in the measurements. In such cases, using a vector basis derived directly from the experimental data can be more justified and less arbitrary than fitting a pre-selected set of basis functions. Thus statistical methods are predicated on the covariance of the measurements, and thereby on any symmetries of the data distribution.

The covariance matrix $\mathbf{C_x}$ of an m-dimensional time series **x**(t) (2.6) is

$$\mathbf{C_x} = \frac{1}{N}\sum_{t=1}^{N}\left(\mathbf{x}(t_i)-\overline{\mathbf{x}}\right)^{\dagger}\left(\mathbf{x}(t_i)-\overline{\mathbf{x}}\right) \qquad (3.8)$$

where the superscript † denotes the transpose of the vector. The diagonal elements in (3.8) are the variances of the variables (Eq. (2.5). Normalizing the covariance with the standard deviations of the variables involved gives the correlation. For instance, the autocorrelation function of a scalar time series x(t) is the normalized covariance of x(t) and x(t+τ):

$$C_x(\tau) = \frac{1}{T}\frac{1}{\sigma_x^2}\int_0^T\left(x(t)-\overline{x}\right)\left(x(t+\tau)-\overline{x}\right)dt \qquad (3.9)$$

where τ is the time lag between two measurements and the variable is a continuous function of time.

Peaks of $C_x(\tau)$ represent periodicities of the plasma system at the linear approximation. Both the autocorrelation and the power spectral density measure the contribution of a mode in the time series x(t), in the time and frequency domain, respectively. The fundamental relation between spectral and statistical methods is exemplified by the Wiener-Khinchin theorem which states that the autocorrelation function is the Fourier transform of the power spectral density:

$$C_x(\tau) = \int_{-\infty}^{\infty}\left|\widetilde{X}(\omega)\right|^2 \mathrm{e}^{i\omega\tau}\, d\omega$$

where the limits of integration of (3.9) have been extended to infinity.

Figs. 11 and 12 show the power spectrum and autocorrelation function of the daily solar wind radial velocity in the equatorial plane, from the OMNI dataset provided by NASA/NSSDC. The solar rotation period and a harmonic and subharmonic are indicated. Different variables of the same system may have different autocorrelation times, which indicates that they are subject to different dynamics. For example, the autocorrelation time of various IMF components is much smaller than that of the solar wind speed (Fig. 13).

Statistical methods of counting degrees of freedom extract this information from the covariance matrix. A frequently used approach is principal components analysis (PCA; also named singular spectrum analysis, Karhunen-Loeve analysis, etc.) [Press et al., 1992; Jackson, 1991; Preisendorfer and Mobley, 1988]. Other approaches include factor analysis and cluster analysis.**Principal component analysis.**

PCA determines a vector basis by diagonalizing the covariance matrix $\mathbf{C}_\mathbf{x}$ (3.8). For simplicity the average activity is zero, $\overline{\mathbf{x}} = 0$. Since $\mathbf{C}_\mathbf{x}$ is positive definite and symmetric, its eigenvalues are positive:

$$\mathbf{C}_\mathbf{x}\mathbf{e}_i = w_i^2\mathbf{e}_i,\ i = 1,2,\ldots,m \tag{3.10}$$

where w_i^2 and $\mathbf{e}_i$ are the i-th eigenvalue-eigenvector pair, and more importantly the m eigenvectors form an orthogonal vector basis.

The original vector time series $\mathbf{x}(t_i)$, i=1,2,…,N (which is an Nxm matrix $\left[\mathbf{x}(t)\right]$) can be projected on the eigenvector basis. The matrix is written:

$$\left[\mathbf{x}(t)\right] = \mathbf{U}\cdot\mathbf{W}\cdot\mathbf{V}^\dagger \tag{3.11}$$

where **V** is an orthonormal matrix whose rows are the eigenvectors $\mathbf{e}_i$, **W** is a diagonal matrix whose non-zero elements are the eigenvalues w_i^2, and orthogonal matrix **U** contains the projection of the time series data. By convention the eigenvalues in **W** are ranked in order of decreasing magnitude. Note that **U** contains the time dependence and is also an Nxm matrix, whereas **U** and **V** are time-independent and mxm. Eq. (3.11) is called the singular value decomposition (SVD) theorem for matrix $\left[\mathbf{x}(t)\right]$, and w_i are called the singular values of the matrix (for the numerical SVD method see [Press et al., 1992]). The decomposition gives m new variables, called the principal components:

$$p_i(t) = \mathbf{x}(t)\cdot\mathbf{e}_i = \mathbf{u}(t)\cdot w_i,\ i = 1,2,\ldots,m \tag{3.12}$$

In fact the eigenvalue w_i^2 is the variance of $p_i(t)$, or of the original data along the direction $\mathbf{e}_i$. The variance is maximized along the direction of $\mathbf{e}_i$ and minimized in other directions. A linear algebra theorem guarantees the uniqueness of this decomposition [Golub and Van Loan].

The most significant principal components are interpreted as the effective degrees of freedom. If the physical processes that contribute to **x**(t) are uncorrelated, the separation of the variance in components may be used to separate and identify the effects of each process at least in the low-order principal components. However, there is significant mixing in the higher-order components.

The reader may recognize the similarities between PCA and minimum variance analysis (MVA), which determines the main magnetic field directions in a plasma structure [Sonnerup and Cahill, 1967; Dunlop, 1995; Samson, 1983]. The experimental time series are fluctuations in the three components $\delta B_i(t)$ of the magnetic field **B** in the plasma frame of reference [Dunlop et al., 1995], and $\left[\mathbf{x}(t)\right]$ is a 3xN matrix. The eigenvectors of the covariance matrix represent the directions in which the field varies most, intermediate, and least. The eigenvectors with the two highest variances provide a tangential approximation of the current sheet structure, while the minimum-variance vector defines the normal direction. In this way MVA is used to estimate features (position, orientation, width, etc.) of 1-D and 2-D current sheets.

In PCA, the effective degrees of freedom are the projections on the significant eigenvectors (cf. (3.12)), where significance of eigenvector $\mathbf{e}_i$ is measured as the magnitude of singular value w_i (or eigenvalue w_i^2). If the noise variance is σ_{noise}^2, significant eigenvectors correspond to singular values

$$w_i > \sigma_{noise};\ i = 1, 2, ..., s$$

and $df_{eff} = s$. The singular value spectrum w_i decays rapidly with i, so typically $s \ll m$.

The projections on the remaining vectors $\mathbf{e}_i$, $s < i \le m$, can be subtracted from the original data $\mathbf{x}$(t), in what is effectively a nonlinear noise-reduction method. The original j-th time series x_j (j=1,2,…,m) is replaced by the model:

$$\hat{x}_j(t) = \sum_{i=1}^{s} u_{ji}(t)\, w_i \mathbf{e}_i \qquad (3.13)$$

Advantages and limitations.

A justification for empirical methods, such as PCA, is a certain objectivity when the dynamics are not known in sufficient detail and there are no clear symmetries in the system geometry or interactions.

In particular, the PCA method determines df_{eff} and the dynamics directly from the time series data rather than from preconceived models. The df_{eff} is often underestimated because singular values w_i decay rapidly with rank i. PCA-based models are therefore of lower order than e.g., Fourier or spherical-harmonic expansions. Any resulting models will therefore be more parsimonious (efficient). Third, PCA is useful for reducing low-amplitude noise in the data: setting the lowest singular values $w_i < w_{noise}$ to zero, and inverting the PCA, essentially filters out the contribution of the low-power degrees of freedom to each variable.

However, the physical interpretation of results from PCA and other statistical methods may be difficult. The meaning of modes may be unclear, especially for multivariate time series $\mathbf{x}$(t). The relative normalization of components of a multivariate series may further obscure the physical meaning of the projection variables p_i(t). In addition, while the decomposition of the variance may be clear for low-order w_i's, the orthonormality condition $\mathbf{e}_i \mathbf{e}_j = \delta_{ij}$ (which guarantees that the p_i(t) are exactly uncorrelated) lead to oscillations in the eigenvectors. The oscillation frequency increases linearly with order i [Gibson et al., 1992]. In contrast to p_i(t), the original variables x_i(t) are seldom completely uncorrelated.

Applications.

High-latitude geomagnetic disturbances were analyzed into two modes by Sun et al. [2000]. Figs. 14 shows the original equivalent current density while Fig. 15 shows the first two principal components (dubbed "natural orthogonal components by Sun et al. [2000]). By correlating the geomagnetic patterns with the solar wind and IMF history, they identified the first mode as primarily directly-driven by the solar wind input (convection-induced currents). The second mode was identified as a loading-unloading response due to the storage of energy in the magnetic lobes and plasma sheet and later release during a substorm. The corresponding current system is the substorm current wedge. Fig. 16 shows the singular values w_i corresponding to the principal components from Eq. (3.10). Note the rapid decline of the singular spectrum which indicates that the first few eigenvalues are sufficient to represent the entire disturbance shown in Fig. 14.

Interesting results for PCA have been obtained from its application to dynamics of high-latitude magnetic indices [e.g., Sharma et al., 1993]. The significant singular values between the AL index and the VB_s component of the interplanetary electric field have been interpreted as directly-driven and loading-unloading [Sitnov et al., 2000].

3.4. The state space approach

Like other statistical methods, state space methods represent the dynamics and estimate the degrees of freedom by introducing state variables whose number is df_{eff}. They are considered separately here because of their significance for systems theory.

For deterministic systems, the estimation of the number of degrees of freedom is linked to the study of state space trajectories. A useful analysis of trajectories is related to different types of equilibria. Equilibria of interest include the stable fixed point, the limit cycle, and the stable torus.

A static equilibrium is called a fixed point, $\mathbf{x}(t)$=const. Trajectories in the vicinity of stable fixed points are "attracted" to them (reach them asymptotically), while trajectories near unstable fixed points are "repelled" away from them. If there are more than one stable fixed points, the part of the state space where trajectories reach one of them, is called that fixed point's basin of attraction. There are no dynamics in such systems and $df_{eff}=0$.

Time-dependent equilibria include periodic trajectories. For instance the Van der Pol oscillator:

$$\frac{d^2x}{dt^2}-\left(\varepsilon-x^2\right)\frac{dx}{dt}+x=0 \tag{3.14}$$

has been used as a model of nonlinear oscillations in lab and space plasmas. It has a stable equilibrium, namely a periodic, nonharmonic trajectory called a limit cycle with size approximately equal to $\sqrt{\varepsilon}$. All trajectories in phase space, with initial conditions either $\left|\mathbf{x}_0\right|<\sqrt{\varepsilon}$ or $\left|\mathbf{x}_0\right|>\sqrt{\varepsilon}$ rapidly reach the limit cycle. At that point only one variable is sufficient to describe the system's time development on the limit cycle, and $df_{eff}=1$.

The driven van der Pol oscillator

$$\frac{d^2x}{dt^2}-\left(\varepsilon-x^2\right)\frac{dx}{dt}+x=F\left(t\right) \tag{3.15}$$

has been used to model thermionic discharge and double plasma devices [Klinger et al., 1995; Klostermann et al., 1997]. In these systems, microscopic potential structures accelerate ions into regions of negative space charge. The parameter ε represents the net nonlinear effect of this acceleration while F(t) represents the external voltage. At auroral latitudes, modulation and spectral broadening of electrostatic ion-cyclotron waves has been attributed to mode conversion effects by currents such as the auroral electrojet [Koepke et al., 1994]. The variation of plasma parameters, represented by F(t), influences the modes that the current can excite such as the ion-cyclotron waves as well as other waves and instabilities [Chaturvedi and Ossakow, 1990]. The df_{eff} can be 1 or 2, depending on the characteristics of F(t).

The generalization of a periodic trajectory for an m-variable system is a torus embedded in the m-dimensional state space. For a stable torus, trajectories are attracted to it and therefore $df_{eff}<m$. In fact, df_{eff} is the dimension of the torus (which is analogous to the limit cycle of system (3.14)). The dimension of the torus is the number of incommensurate frequencies in the system dynamics (if two or more frequencies are commensurate, only one of them is counted as a degree of freedom).

The stable fixed point, limit cycle, and stable torus are examples of attractors, structures in the state space where the system eventually is confined to. Complex attractors are produced by non-periodic dynamics (cf. Eq. (3.15) and the effects of F(t)), and nonlinear, or chaotic, dynamics (discussed in Sec. 5.3 below).

State space methods that estimate df_{eff} are designed to measure the dimension of the attractor. Typically the attractor is visualized through a trajectory $\mathbf{x}(t)$, which is a vector time series either directly measured or reconstructed. The dimension of the trajectory itself is not of interest: at best it equals 1 if the trajectory is continuous;[6] at worst it biases the df_{eff} estimate. Therefore df_{eff} is calculated from the intersection of $\mathbf{x}(t)$ with a fixed surface in state space, called the Poincaré surface of section and chosen to

[6] Discrete-time systems have discontinuous trajectories, generally of non-integer dimension.

intersect the attractor. After the degrees of freedom are calculated, we need to add 1 (if **x**(t) is continuous) in order to obtain df_{eff}.

The intersection of the trajectory with the Poincaré surface of section forms a fractal set. The fractal scaling of the intersection can be calculated with the capacity dimension d_0, Eq. (2.3). If the number N of intersection points is low, a different scaling function can be used which relies on pairs of points and improves the statistics. The function is the correlation integral and it counts the pairs of points $\left(\mathbf{x}_i, \mathbf{x}_j\right)$ whose distance is less than r:

$$CI_{\mathbf{x}}(r) = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j \neq i} \Theta\left(\left|\mathbf{x}_j - \mathbf{x}_i\right| < r\right) \tag{3.16}$$

where Θ is the Heaviside (or step) function taking only the values 0 or 1. The normalization ensures that for large r, $\lim_{r \to \Delta \mathbf{x}_{\max}} CI_{\mathbf{x}}(r) \to 1$. If the correlation integral scales with r as

$$CI_{\mathbf{x}}(r) = c \cdot r^{d_2} \tag{3.17}$$

where c is a constant, then the correlation dimension d_2 is

$$d_2 = \frac{\ln CI_{\mathbf{x}}(r_2) - \ln CI_{\mathbf{x}}(r_1)}{\ln r_2 - \ln r_1} \tag{3.18}$$

or the slope of $C_{\mathbf{x}}(r)$ in a double-logarithmic plot (cf. with Eq. (2.3)).For the Langmuir oscillation system, Eq. (3.6), one would apply an embedding method (Sec. 2.3) to the observed time series $n_1(t)$ and construct state vectors **x**(t). Then one would measure the fractal dimension d_0 or d_2 to estimate df_{eff}. More complex estimates include the multifractal spectrum whose elements include d_0 and d_2 .

A different type of df_{eff} estimate is obtained as the order of a dynamical model built to reproduce the time series data **x**(t). If the model is highly accurate in representing the dynamics, the order can be interpreted as df_{eff}. We will see such examples in Sections 4 and 5.

Advantages and limitations.

State space methods are more flexible compared to other statistical methods that work directly with the covariance matrix [Abarbanel et al., 1993]. Part of this flexibility is due to the geometric properties of methods which preserves characteristics of the probability density and makes them available for visualization. Another advantage is due to the range of the embedding parameters as well as the dimension-estimation parameters.

However, fractal-analysis methods also have pitfalls the most important of which are the following:

- In the ideal case, noise-free, stationary, long time-series data should be used. Noise can distort trajectories and produce artificial intersections which appear as false returns to equilibria.
- Short time series have been the source for many invalid low-dimension claims. A conservative rule of thumb is that the dimension estimate that exceeds $\log_{10}N$ is suspect. Methods that rely on pairs of points should be used as in (3.16).
- Oversampling tends to blur the distinction between low- and high-order dynamics. For example, estimation of d_2 is based on the local density (Eq. 2.2) and in principle is independent of the time resolution. If the resolution is too high, however, the density along the trajectory is much higher than in the transverse directions, the ones which are of interest (Fig. 17). This can lead to an artificial decrease of the correlation dimension [cf. Prichard and Price, 1992]. Oversampling is corrected by decimating the time series or, in embedding methods, by increasing the delay time to $\tau \gg \Delta t$ [Theiler, 1986].
- If the state space dimension is too low, trajectories contain artificial intersections and turning points. Varying the dimension removes these false intersections and separates false neighboring points. An optimal state space dimension is estimated by measuring how rapidly false intersections and nearest neighbors decay as the dimension is increased [Kennel et al., 1992]. An example of the dependence on the state space dimension, taken from a nonlinear modeling approach called neural networks [Gleisner and Lundstedt, 2001], is shown in Fig. 18.

- Stochastic time series with a rapidly declining power-law power spectrum ("colored random noise") produce a low correlation dimension d_2 which however is a consequence of the signal's long autocorrelation rather than any deterministic dynamics [Osborne and Provenzale, 1989; Higuchi, 1990]. The dimension d_2 is related to the spectral index and not to the degrees of freedom. Smoothing the data can suppress low levels of this type of noise.
- Nonstationarity can be a serious complication for fractal analysis and lead to misestimation of dimension. Still there are mitigating schemes [Theiler; Prichard and Price, 1992].

Applications.

In the 1990s a debate focused on whether the large-scale dynamics of the magnetosphere can be attributed to a small df_{eff} representing a small number of strongly interacting modes. The new elements were nonlinearity and low-order (low dimensionality) put forward to account for the limited success of earlier oscillator-type models, electronic circuit analogues, and data-based filters. Several studies examined correlation dimension of auroral electrojet geomagnetic index time series [Vassiliadis et al., 1990; Roberts et al., 1991; Shan et al., 1991; Pavlos et al., 1992]. The AL and AE indices were chosen since they represent the geomagnetic activity due to regional current systems determined by high-latitude convection and substorms. State spaces were created using embedding and the correlation dimension d_2 was found to be low, in a range of 2.5-4.0 that varied with the interval studied.

However, the correlation dimension results were criticized because nonstationarity of the index time series led to underestimates of d_2 [Prichard and Price 1992, 1993]. Moreover, the nonstationarity is primarily due to the solar wind driver rather than internal magnetospheric dynamics, and therefore the solar wind activity level must be somehow accounted for in the calculation [Price and Prichard, 1993, 1994]. No fractal dimension estimators exist for input-output systems.

Nevertheless the most recent studies have found statistically significant evidence for the correlation dimension method addressing the nonstationarity issue [Pavlos et al., 1999a,b]. Also while incorporating the solar wind information is necessary, the df_{eff} can be estimated from input-output ARMA models (Sections 3.4, 5.4, and 5.6) rather than from fractal dimension estimators.

In summary, although the issue of the nonlinear response is settled, the number of degrees of freedom is still debated. Approaches such as self-organized criticality provide evidence for a multiscale (i.e., many-degree-of-freedom) dynamics [e.g., Lui et al., 2000; Uritsky et al., 2002]. Therefore it is probably more relevant to differentiate intervals in which the magnetosphere dynamics is of high order and stochastic from those when it becomes low-order and coherent, and to identify the physics of the transition between the two.

4. Linear dynamics and response

4.1. Feedback and dissipation

Estimating df_{eff} is a first step in estimating a plasma system's complexity, but determining the dynamics of these degrees of freedom gives an insight into the physical processes controlling the system. The question here is: what is the simplest linear dynamics that explains the short-term behavior of the system? "Short-term" is in reference to internal time scales of the system, and especially the lifetime of the present equilibrium state. In order to place the answers in a physical context, it is important to first review feedback mechanisms.

Plasmas exhibit positive feedback through the development of instabilities (Fig. 19a) which lead to an exponential growth in energy and a decrease in entropy. On the other hand, negative feedback or dissipation of energy is possible through wave-particle and wave-wave interactions as well as collisions. In the Langmuir oscillation system of Sec. 3.1, the ion drag term of Eq. (3.2) provides an instance of negative feedback causing the oscillation amplitude to decay with time.

Example: two-stream instability. Returning to the plasma oscillations (3.1), if the electrons stream past the ions with velocity $v_0 \neq 0$ the two-stream instability is excited [Chen, 1984; Sturrock, 1994]. This fundamental kinetic instability occurs in several geospace plasma environments including field-aligned currents in the auroral region; the tail-disruption region on the nightside of Earth (10-12 R_E) [Lui et al., 1996], the interplanetary medium where the oscillating two-stream instability is excited; and the equatorial ionosphere where the Farley-Buneman instability occurs [Oppenheim et al., 1995].

For a finite velocity v_0, the electron continuity and momentum equations are:

$$\begin{aligned} &\frac{\partial n_{e1}}{\partial t} + n_0 \frac{\partial v_{e1}}{\partial x} + v_0 \frac{\partial n_{e1}}{\partial x} = 0 \\ &m_e n_0 \left[\frac{\partial v_{e1}}{\partial t} + (\mathbf{v}_0 \cdot \nabla) v_{e1} \right] = -e n_0 E \end{aligned} \tag{4.1}$$

The corresponding equations for the ions (which cannot be assumed immobile) are:

$$\begin{aligned} &\frac{\partial n_{i1}}{\partial t} + n_{i0} \frac{\partial v_{i1}}{\partial x} = 0 \\ &m_i n_0 \frac{\partial v_{i1}}{\partial t} = e n_0 E \end{aligned} \tag{4.2}$$

The electric field is specified from Poisson's equation:

$$\nabla E = \frac{e}{\varepsilon_0}\left(n_{i1} - n_{e1}\right) \tag{4.3}$$

With an ansatz of exp(ikx) for the spatial dependence of all variables, the result is two coupled equations for the density fluctuations:

$$\frac{\partial^2 n_{e1}}{\partial t^2} + 2ikv_0 \frac{\partial n_{e1}}{\partial t} - k^2 v_0^2 n_{e1} = \frac{n_0 e^2}{m_e \varepsilon_0}\left(n_{i1} - n_{e1}\right) \tag{4.4}$$

$$\frac{\partial^2 n_{i1}}{\partial t^2} = -\frac{n_0 e^2}{m_i \varepsilon_0}\left(n_{i1} - n_{e1}\right) \tag{4.5}$$

Forward modeling.

If the equations (4.4) are known a priori, a scaling exp(iωt) of the densities with a complex frequency $\omega = \omega_r + i\omega_i$, leads to the dispersion relation

$$1 = \omega_p^2 \left[\frac{m_e / m_i}{\omega^2} + \frac{1}{\left(\omega - kv_0\right)^2} \right] \tag{4.6}$$

The frequency is Doppler-shifted by the electron velocity v_0. For $kv_0 \approx \omega_p$, the growth rate of the instability, or the imaginary part of the frequency, is

$$\omega_i \simeq \omega_p \left(\frac{m_e}{m_i} \right)^3 \tag{4.7}$$

which is the rate at which energy increases. The main structure in the state space is a stationary or slowly traveling ($\omega_r \ll \omega_i$) wave of wavelength equal to $2\pi v_0 / \omega_p$ (Fig. 19b,c). Eventually, the linear growth slows down and nonlinear effects appear such as particle trapping (Fig. 19d).

The occurrence of the instability depends on the magnitude of v_0. For $kv_0 \gg \omega_p$ the instability is not excited, and the dynamics must be modeled at microscopic scales (e.g. through numerical simulation). For $kv_0 \approx \omega_p$, however, the instability is excited, and the dynamics is accounted for by four variables in Eq. (4.4), for example (n_{e1}, dn_{e1}/dt, n_{i1}, dn_{i1}/dt). Other combinations of the densities, electric field and velocity are possible. If the variables are chosen properly, two grow exponentially at a rate (4.7) and the two complex conjugates decay and can be neglected. Thus the system is represented by two degrees of freedom in the linear growth regime.

In this way feedback mechanisms reduce the number of effective degrees of freedom (cf. discussion in Section 2.1). The low-order approximation usually breaks down when the instability saturates.

Inverse modeling.

If the dynamics is not known a priori, but there are experimental data (time series) of the electron density, we can reconstruct equations analogous to (4.4) by using state space methods as discussed below.

4.2. State space methods

There is a large bibliography on development of state space models and other empirical methods to reproduce nonlinear systems [Casdagli and Eubank, 1992; Abarbanel et al., 1993; Weigend and Gershenfeld, 1994]. Continuing from Section 2.3, the procedure of developing a state space from time series measurements is called embedding, or state space reconstruction. Embedding theorems show that for sufficiently large datasets and low levels of noise, the new variables are equivalent to the original ones [Takens, 1985; Sauer et al., 1991]. "Equivalent" means that the topological properties of the system (geometrical features in the state space, such as those associated with an equilibrium, periodic trajectory, or stability property) are invariant from one space to the other. Therefore estimating statistical and dynamical properties of the system, such as df_{eff} or the instability's growth rate, gives the same result whether calculated in the reconstructed or the original state space [Packard et al., 1980].

The most common methods for reconstructing a state space are based on time derivatives, time delays, and empirical orthogonal functions such as principal components [Abarbanel et al., 1993].

a. **State spaces based on time derivatives.** This type of state space is the most familiar. Consider the Langmuir oscillation system (3.1). The linearized equations can be written as:

$$\frac{d\mathbf{x}}{dt} = \mathbf{A}\cdot\mathbf{x} \tag{4.8}$$

where the state vector is $\mathbf{x}(t) \equiv (n_1, dn_1/dt)$ and **A** is an mxm matrix. All variables are expressed in terms of n_1 and its time derivatives.

More generally, in a system described with m differential equations, the state vector at time t contains derivatives of order up to m-1.

$$\mathbf{x}(t) = \left[n_1(t), dn_1/dt, d^2n_1/dt^2, \ldots, d^{m-1}n_1/dt^{m-1}\right] \tag{4.9}$$

According to the inverse-function theorem, any system variable can be expressed in terms of the components of **x**(t) except in the vicinity of $|\mathbf{x}| = 0$. Thus an m-order differential equation for n_1 represents up to m independent modes.

In the Langmuir oscillation of Eq. (3.5), one needs to include up to the first derivative in (4.9): stopping at that order produces the same state space, (n_1, dn_1/dt), as obtained from the equations (Fig. 20a,b). Even if the equations are not known, one can quickly verify that the dynamics of $n_1(t)$ can be written as a function of only two time derivatives (including the zeroth derivative, i.e., $n_1(t)$).

ODE models. A linear model based on (4.9) is an m-order ordinary differential equation (ODE):

$$\frac{d^m \hat{n}_1(t)}{dt^m} = \sum_{i=0}^{m-1} \alpha_i \frac{d^i \hat{n}_1(t)}{dt^i} + \varepsilon(t) \tag{4.10}$$

Here $\widehat{n_1}$ is the model density, and coefficients α_i are time-invariant. A model fitted to data is considered accurate if the fitting error e(t) is significantly smaller than the variance of $n_1(t)$. Model predictions and testing are discussed in a Sec. 4.4.

Comparing with the Langmuir oscillations system, Eq. (3.5), the coefficients of (4.10) are

$$\alpha_0 = -\frac{n_0 e^2}{\varepsilon_0 m_e}$$
$$\alpha_1 = -\frac{\gamma}{m_e} \quad (4.11)$$
$$\alpha_i = 0,\ i > 1$$

Based on the α_i, one can derive the oscillation frequency $\omega_r = \pm\sqrt{\alpha_0}$ and damping rate $\omega_i = \alpha_1$.

Advantages and limitations.

The first few time derivatives can have a physical interpretation. Thus dn_1/dt is often related to a rate of change or to a velocity by using a continuity equation (cf. Eq. (3.1)) while the second derivative is related to an oscillation. Furthermore, derivatives have low correlations with each other and can be used to approximate independent variables. Finally, if the data have sufficiently low levels of noise, ODE models can reproduce a significant amount of the variance of the original time series.

In real-world datasets, however, the number of derivatives needed for an accurate representation is typically high. Derivatives amplify signal distortions due to instrument noise, and this problem increases with the derivative order. The above effects are significant obstacles in using derivatives to model actual data.

b. **Time delays.**

Much more commonly used than derivatives are time delays, or lags, $x(t-\tau)$ produced by timeshifting the data by the delay time τ a number of times (cf. Sec. 2.3). For simplicity it is assumed that only one variable is observed. For the Langmuir system, the delay vector is

$$\mathbf{x}(t) = \left[n_1(t), n_1(t-\tau), n_1(t-2\tau), \ldots, n_1(t-(m-1)\tau)\right] \quad (4.12)$$

Fig. 20c shows the components of a two-dimensional state vector plotted against each other.

The delay, or lag, τ is a free parameter and represents a periodic sampling of the plasma activity. Its default value is the time resolution Δt. To increase model efficiency, however, delay coordinates should have little correlation as possible (ideally they should be independent). To accomplish that, τ is determined as the smallest time that satisfies

$$\frac{\partial C_{n_1}}{\partial \tau} = 0, \frac{\partial^2 C_{n_1}}{\partial \tau^2} > 0 \quad (4.13)$$

or

$$C_{n_1}(\tau) = 0$$

where $C_{n_1}(\tau)$ is the autocorrelation function (3.9) for $n_1(t)$ (see also [Abarbanel et al., 1993]. Instead of the autocorrelation function the mutual information can be used [Fraser and Swinney, 1986]. The mutual information measures for nonlinear correlations that the autocorrelation functions ignores.

The number m of delays used is called the embedding dimension, because the observed time series data are embedded in the m-dimensional state space.

AR models. The model built from delay variables of the electron density $n_1(t)$ has the form:

$$\hat{n}_1(t+1) = \sum_{i=0}^{m} A_i \hat{n}_1(t - i\tau) = \mathbf{A} \cdot \mathbf{x}(t) \quad (4.14)$$

where vector **A** contains all coefficients a_i. Model (4.14) is called autoregressive (AR), because it regresses the density fluctuation at t+1 on recent fluctuations. The order m of the optimal ARMA model is an estimate for df_{eff}.

The plasma oscillation equation, (3.5), can be discretized in time and compared to (4.14) with $\tau = 1\ \Delta t$. The comparison yields:

$$a_0 = 2 - \omega_L^2 \Delta t^2 + \frac{\gamma}{m_e}$$
$$a_1 = -1 - \frac{\gamma}{m_e} \qquad (4.15)$$

while higher order a_i terms are zero. In practice, model (4.14) is directly fitted to the data.

From (4.15) the time dependence is an oscillation $e^{i(\omega_r + i\omega_i)t}$ with complex frequency

$$\omega_r = \pm \frac{\sqrt{1 - a_0 - a_1}}{\Delta t}, \quad \omega_i = \frac{1 + a_1}{\Delta t} \qquad (4.16)$$

Advantages and limitations.

Delay coordinates are preferable to derivatives for several reasons. First, they have a higher tolerance of measurement noise. Second, they are more general: discrete-time derivatives can be approximated as linear combinations of delay coordinates, but the reverse is not always possible. (In some cases, however, delay models are approximated by less-accurate ODEs, which are more amenable to physical interpretation). In the two-stream instability (4.4), one can in principle regress a fourth-order ODE model (4.10) to the measured data. Instead, typically one fits a fourth-order AR model (4.14) to the time series data, because AR models are more flexible than ODEs in handling noise and discontinuities in the signal. Then, to interpret the coefficients physically, a fourth-order ODE might be discretized in time and compared to the AR model. Because of these properties, delays are the variables of a large variety of model types (time series models, neural networks, etc.).

There are certain limitations to the applicability of delays, however. First, the value of t needs to be specified as one additional free parameter. The default choice is τ=Δt, the time resolution of the data, but that may lead to oversampling as discussed in Section 3.4. Increasing τ reduces oversampling and decreases the dimension m of the delay vector if the time interval length (m-1)τ remains constant. A lower dimension reduces the omputational cost and the complexity of th state space. Second, the physical interpretation of delays is not as clear as in the case of derivatives. A simple interpretation is that, plotted as a waveform, the vector **x**(t) contains the activity x(t) in the time window [t-(m-1)τ, t].

In summary, limitations of delay embeddings are not significant compared to their usefulness as modeling tools.

c. Principal components.

As we saw in Sec. 3.2, for multivariate time series **X**(t) from a linear system, the covariance matrix $\mathbf{C_x}$ (Eq. 3.5) contains key information about the system's activity (again it is assumed that $\overline{\mathbf{x}} = 0$). Eigenvectors $\mathbf{e}_i$ of $\mathbf{C_x}$, with eigenvalues w_i^2, are useful choices as state-space variables.

A PCA model for the fluctuations $\hat{n}_1(t)$ takes the form:

$$\hat{n}_1(t) = \sum_{i=0}^{m} u_{1i}(t) w_i \mathbf{e}_i \qquad (4.17)$$

(cf. Eq. (3.13)). Principal components can, by construction, explain an arbitrarily high percentage of the variance in the measured time series. If only a scalar time series is available, a state representation **x**(t) can be obtained from time-delay embedding (4.12), and the principal components are obtained from the covariance matrix of $\left[\mathbf{x}(t)\right]$.

Advantages and limitations.

Like delay-based models, PCA models have several free parameters that provide flexibility. The typical limitation is the lack of a physical interpretation, which becomes more severe for higher-order eigenvectors.

Summary.

There are several approaches to constructing a state space from time series measurements and they lead to distinct types of models. Among those, AR models express the linear dynamics of a plasma system in the most flexible manner. ODEs are useful for the physical insights they provide, although they are less capable of representing noisy observations as well as AR or PCA models do.

4.3. Interacting plasmas

The systems discussed up to this point are autonomous (closed), and external effects are negligible, except for constant driving or an initial impulsive perturbation. Geospace plasmas, however, continuously exchange mass, momentum, and energy with their environment. It is therefore important to study a system of two or more interacting plasmas.

To keep things simple, consider two adjacent plasmas that exchange energy and momentum. The interaction is asymmetric: one of the plasmas (the "environment") exerts a stronger effect on the other (the "system" of interest) than vice versa. Examples include a stellar wind and a planetary magnetosphere, a magnetosphere and ionosphere, an accretion disk and a central star, , etc. It is assumed that environment, or input, variables **u**(t) determine the system, or output, variables **x**(t), although some of the variables are not directly measured.

Continuing with the electron oscillations of Eq. (3.1), consider that they are driven by a time-dependent external electric field, $E_{ext}(t)$. Again ion motion is neglected, except its effects on the electrons as the frictional term $-\gamma n_0 v_1$. As before, the observable output is the electron density perturbation $n_1(t)$. The equations are:

$$\begin{aligned} &\frac{\partial n_1}{\partial t} + n_0 \frac{\partial v_1}{\partial x} = 0 \\ &m_e n_0 \frac{\partial v_1}{\partial t} = -e n_0 E - \gamma n_0 v_1 + E_{ext} \\ &\frac{\partial E}{\partial x} = -\frac{e}{\varepsilon_0} n_1 \end{aligned} \tag{4.18}$$

where the electric field E(t) is a random variable specified only by its statistical moments, and is independent of position.

The strength of the coupling can be quantified in various ways, a most basic one being the correlation coefficient between input and output:

$$C_{E_{ext},n_1} = \frac{1}{T}\frac{1}{\sigma_{E_{ext}}\sigma_{n_1}} \int_0^T \left(E_{ext}(t) - \bar{E}_{ext}\right)\left(n_1(t) - \bar{n}_1\right) dt \tag{4.19}$$

where standard deviations $\sigma_{E_{ext}}$ and σ_{n_1} are used to normalize the covariance. A value c of the correlation coefficient is interpreted as saying that the electric field determines c^2 of the variance of $n_1(t)$, provided the prediction error is uncorrelated with the electron density,

$$\sum_{t=1}^{T} \left(\hat{n}_1(t) - n_1(t)\right) n_1(t) = 0 \tag{4.20}$$

Typically the response to the electric field is not instantaneous, but has a finite time delay. The cross-correlation function:

$$C_{E_{ext},n_1}(\tau) = \frac{1}{T}\frac{1}{\sigma_E \sigma_{n_1}} \sum_{t=1}^{T} \left(E(t) - \bar{E}\right)\left(n_1(t+\tau) - \bar{n}_1\right) \tag{4.21}$$

(cf. Eq. 3.8). For τ=0, one recovers the correlation coefficient (4.19). The time to maximum τ_{max} in C_{E_{ext},n_1}, found from $\frac{\partial C_{E,n_1}}{\partial \tau} = 0,\ \frac{\partial^2 C_{E,n_1}}{\partial \tau^2} < 0$, is an estimate for the time of maximum linear coupling, or the "sluggishness" of the system response. Other local maxima in C_{E_{ext},n_1} represent either instances of secondary coupling between the two systems, or periodicities of the systems.

Example: Particle acceleration in the inner magnetosphere. Solar wind activity, in particular velocity variations, lead to energization of the inner magnetosphere and electron acceleration in the radiation belts. The time scale of interaction was measured via the correlation between relativistic electron flux (the output of the acceleration process) and solar wind speed (an important input) [Paulikas and Blake, 1979]. Scatterplots of the flux versus preceding solar wind velocity at different lags showed that the peak response occurs 2-3 days after the solar wind speed begins to increase. The long time scale indicated that the interaction involves several stages including excitation of low-frequency waves (as demonstrated later). Comparisons with other interplanetary input variables such as solar wind density and IMF components showed that solar wind velocity had the highest correlation amplitude, and was well above that corresponding to the IMF. The results indicated that acceleration processes were the result of viscous interactions, which are driven by the velocity, much more than of magnetic reconnection, which is driven by the IMF.

The correlation function (4.21) is related to a simple linear regression model. For the electron density one can write

$$n_1(t) = aE(t - \tau_{\max C}) + b \tag{4.22}$$

where $\tau_{\max C}$ is the peak time of the cross-correlation function.

Finite-Impulse-Response Models.

The cross-correlation function assumes an instantaneous, time-shifted response. However, the response of a realistic plasma system consists of several stages and as a result has a finite duration T. The present state **x**(t), represented by a scalar output x(t), is determined by the input activity over the recent time [t-T,0], i.e. the past inputs u(t), u(t-1), …, u(t-T) (Fig. 21). The width of the correlation function is an indication of the magnitude of T. The output can be written as a weighted average of the past inputs:

$$x(t) = x(t-T) + \int_0^T H(\tau) I(t-\tau)\, d\tau \tag{4.23}$$

It is often assumed that x(t-T)=0 (zero-state). Referring to the electron plasma system (4.18) a density perturbation at t is due to several recent electric field inputs:

$$\hat{n}_1(t) = \int_0^T H(\tau) E_{ext}(t-\tau)\, d\tau$$

Eq. (4.23) is the finite impulse response (FIR) model (formerly "moving-average" model). H(t) is called the impulse response function (also propagator or Green's function) and represents the most general, linear, time-invariant coupling between x(t) and u(t). The time T represents the memory time of the system relative to external perturbations. This time is finite, hence the name of the response. If the electric field is an impulse, $E_{ext}(t) = E \cdot \delta(t)$, the response is simply

$$\hat{n}_1(t) = H(t), \qquad 0 < t < T$$

and zero at any other time. This is the physical meaning of the impulse response function.

Forward modeling.

Assuming a harmonic time dependence $n_1 \sim \exp(i\omega t)$, one can derive the impulse response function as a function of the oscillation frequencies of the system. For Eq. (4.18) the impulse response is

$$H(\tau) = \frac{e^{-\omega_1 \tau} - e^{-\omega_2 \tau}}{\omega_2 - \omega_1} \tag{4.24}$$

where

$$\omega_{1,2} = \omega_p \pm i\gamma$$

Inverse modeling.

If only experimental data are available, the impulse response function can be obtained from by fitting (4.23) to them. For simplicity it is assumed that the time resolution Δt=1. A classical numerical method involves multiplying with the transpose of $E_{ext}(t)$ thus forming the covariance matrix of the input and the cross-correlation matrix between input and output [Clauer, 1986]. The formal solution for H is then

$$\mathbf{H} = \underbrace{\left[\left(E_{ext}(t)\right)^T E_{ext}(t)\right]^{-1}}_{\text{Autocovariance}} \underbrace{\left(\left(n_1(t)\right)^T E_{ext}(t)\right)}_{\text{Cross-covariance}} \tag{4.25}$$

where $\mathbf{H}(\tau)$ is a T-dimensional vector. Algebraic-system solvers are used to determine H either from (4.25) or (4.23) [Press et al., 1992; Vassiliadis et al., 1995]. Typically the system of linear equations is overdetermined (in fact $T \ll N$, where N is the number of time series data) and leads to singularities in the covariance matrix. Techniques such as singular value decomposition (SVD) [Press et al., 1992; Jackson, 1991] and other methods are used to regularize singular solutions. Some of these methods also estimate the uncertainty ΔH_i in the model coefficients.

FIR models are powerful tools for understanding interacting plasmas. They provide more detailed coupling information than the cross-correlation function. Although the impulse response may not always reproduce the observed dynamics in detail, it is a reliable modeling tool. Coefficients of models based on the impulse response are generally more robust than some specialized nonlinear and adaptive techniques. In addition, the following information is obtained:

- The peaks of H(τ) indicate the time and amplitude of the strongest linear coupling. An example is a work on the geomagnetic index D_{st}, which introduced FIR modeling ("linear prediction filtering") in space plasma physics [Iyemori et al., 1979]. The study described the development of an FIR model and its use in representing the D_{st} dynamics, an approximate measure of the geomagnetic effects of the ring current at midlatitudes. The current's predominantly ion population responds to the interplanetary input $V_{SW}B_{South}$ where V_{SW} is the solar wind speed and B_{South} is the Southward (antiparallel to Earth's dipole) component of the interplanetary field. The product $V_{SW}B_{South}$ is thus an approximate measure of the rate at which the magnetic field is available for reconnection at the magnetospheric dayside. The Iyemori et al. [1979] study showed that the main peak of the average response occurred several hours after the solar wind impact, and corresponds to the speed of convection of plasma sheet ions that feed into the ring current. The H(τ) waveform also provided the average durations and profiles of the storm's three phases (commencement, main phase, and recovery). Iyemori et al. [1979] and others used the calculated response from one interval to predict others.
- Similar to the peaks of the cross-correlation function, the time lag of the largest |H(t)| indicates the delay of the response relative to a change in the input.

After the Iyemori et al. [1979] work on the stormtime ground magnetic response, Nagai [1988] obtained the impulse response of the relativistic electron flux j_e at 2 MeV or higher to the K_p geomagnetic index. The peak response occurred at τ=2 days, in agreement with earlier correlation results [Paulikas and Blake, 1979]. In addition, H(τ) revealed the time profile of the response following the solar wind arrival at the magnetospheric boundary. Subsequent studies revealed a second, negative peak at τ=-1 day, showing that an electron storm is preceded by a temporary decrease in electron flux less than one day earlier [Baker et al., 1990]. The decrease occurs due to the strengthening of the ring current intensity, which distorts the electron trajectories and displaces the particles both Earthward and anti-Earthward from the ring current altitude of L=5.

- FIR models can be calculated in the frequency domain. The Fourier transform of (4.23)

$$\tilde{n}_1(\omega) = \tilde{H}_{n_1}^{E_{ext}}(\omega)\tilde{E}_{ext}(\omega) \tag{4.26}$$

expresses the frequency content of the output as a simple function of the frequency content of the input and the transfer function $\widetilde{H}(\omega)$. The transfer function is the Fourier transform of the impulse response function $H(\tau)$.

The coupling between the solar wind and magnetosphere is nonlinear, the linear approximation (4.26) to the relation between the solar wind input B_{South} and the AE geomagnetic index is still of useful value [Tsurutani et al., 1990]. The power spectral density of AE, $\left|\widetilde{AE}(\omega)\right|$ has two power-law regimes with spectral indices –1.02 and –2.42 separated by a spectral break at 5 hours. In contrast, $\left|\widetilde{B}_z(\omega)\right|$ has a power law of –1.42 (Fig. 22a and b). The transfer function between B_{South} and AE therefore has also two power-law regimes, $\left|\widetilde{H}^{AE}_{B_{South}}(\omega)\right| \sim const.$ and $\left|\widetilde{H}^{AE}_{B_{South}}(\omega)\right| \sim \omega^{-0.5}$, with the same spectral break as AE (Fig. 22c). The similarity of the power law exponents for B_{South} and AE at low ω suggests that long-term solar wind variations are transmitted without change through the magnetosphere. At frequencies higher than $(5 hours)^{-1}$, however, the AE spectral density decreases faster than that of B_{South} and suggests an effective dissipation of the power incident on the magnetosphere, probably in the plasma sheet and the ionosphere during substorm intervals.

Autoregressive moving-average models

In addition to external driving, plasma dynamics includes the effects of internal processes. The effects can be represented by combining (4.14) and (4.23) in an autoregressive moving-average (ARMA) model:

$$\hat{n}_1(t) = \sum_{i=0}^{m-1} A_i \hat{n}_1(t - i\Delta t) + \sum_{i=0}^{l-1} B_i E_{ext}(t - i\Delta t) \qquad (4.27)$$

The ARMA model is analogous to, but is more general than, the inhomogeneous ODE

$$\frac{d^m \hat{n}_1(t)}{dt^m} = \sum_{i=0}^{m-1} \alpha_i \frac{d^i \hat{n}_1(t)}{dt^i} + \sum_{i=0}^{m-1} \beta_i \frac{d^i \hat{E}_{ext}(t)}{dt^i} \qquad (4.28)$$

The ODE (4.28) is solved through a Fourier or a Laplace transform depending on the type of initial conditions. To solve Eq. (4.27) we introduce the z transform which is the discrete-time analogue of the Fourier transform

$$\tilde{N}_1(z) = \sum_{t=-\infty}^{\infty} n_1(t) z^t$$

Then the solution of (4.27) is

$$\begin{aligned}\tilde{N}_1(z) &= A(z)\tilde{N}_1(z) + B(z)\tilde{E}_{ext}(z) = \\ &= \frac{B(z)}{1 - A(z)} \tilde{E}_{ext}(z)\end{aligned} \qquad (4.29)$$

and the density can be obtained by the inverse z transform by expanding $\tilde{N}_1(z)$ [Stanford, 1998].

Linear ARMA and inhomogeneous ODE models can provide a good approximation to dynamics of indices such as D_{st} [Burton et al., 1975] and AU. More complex dynamics, such as those of the AL index, are best captured with nonlinear ARMA, neural networks, and Lotka-Volterra series.

4.4. Prediction, model optimization, validation

After a model has gone through the development stage, its accuracy and reliability are tested. Validation compares the model predictions against physical data.[7] Prediction (or reproduction) of data that have been part of the model's training is called in-sample prediction, and may be used in an initial testing phase. Prediction of new data, never presented to the model during training, is called out-of-sample, and is of direct interest in validation. In complex models, the cycle of development, testing, optimization, and validation is repeated several times.

Prediction accuracy is measured by the prediction error, or difference between model prediction and observation at time t:

$$e(t) = \hat{x}(t) - x(t) \tag{4.30}$$

The time rate of growth of e(t) determines the predictive capability of the model. If we start the model from observed data, then the initial condition is $e(0)=0$. From then on, the growth is linear if $\hat{x}(t)$ comes from a linear FIR or ARMA model, and the growth rate is estimated as

$$\frac{e(T) - e(0)}{T} \tag{4.31}$$

For nonlinear models, the growth can be exponential (Sec. 5).

The prediction error is normalized with the standard deviation of x(t):

$$e_{norm}(t) = \frac{e(t)}{\sigma_x}$$

Errors are averaged in different ways. The root-mean-square error

$$e_{rms} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{x}_i - x_i\right)^2} \tag{4.32}$$

[Bevington and Robinson, 1992] is used as the averaged cumulative error for a number, N, of steps. The rms error is the basis for two oft-used statistics. Normalizing with the standard deviation gives the average relative variance, or the percentage of the data variance not represented by the model predictions

$$ARV = \frac{e_{rms}^2}{\sigma_x^2}$$

The prediction efficiency, $PE = 1 - ARV$, represents the percentage of the variance actually predicted.

A different type of statistic is the correlation coefficient, or data-model correlation:

$$C_{\hat{x},x} = \frac{1}{T}\frac{1}{\sigma_{\hat{x}_1}\sigma_{x_1}}\sum_{t=1}^{T}\left(\hat{x}_1(t) - \bar{\hat{x}}_1\right)\left(x_1(t) - \bar{x}_1\right) \tag{4.33}$$

It is a more robust choice than the prediction error, and this is especially valuable when x(t) is noisy. However, the prediction error gives more detailed information. The two statistics are not independent: if the original signal x(t) is uncorrelated with the prediction (cf. Eq. (4.20)), the prediction efficiency becomes $PE = C_{\hat{x},x}^2$.

Predictions are useful in optimizing a given model and ranking different models depending on how well they reproduce a dataset. Clearly the prediction error depends on all model parameters P that determine $\hat{x}$.

$$e(t;\mathbf{P}) = \hat{x}(t;\mathbf{P}) - x(t) \tag{4.34}$$

In order to optimize the model one seeks to minimize a cumulative prediction error such as e_{rms} or the ARV by changing the parameters. The optimization is formally expressed as

[7] In verification, a concept related to validation, the model is checked for internal consistency and according to specification, and it takes place prior to validation. Verification will not be discussed here

$$\frac{\partial e_{rms}\left(\mathbf{P}\right)}{\partial \mathbf{P}} = 0$$

Optimizing with respect to state space parameters (such as the embedding dimension m) resolves the ambiguity in model specification afforded by embedding theorems (section 2.2). An example of a two-dimensional optimization of an ARMA model for the AL geomagnetic index with an input of $V_{SW}B_{South}$ is shown in Fig. 23. The rms error e(NN,m) is a function of the degree of nonlinearity (NN, in a logarithmic scale) and embedding dimension (m). The position of the minimum shows that the best model is strongly nonlinear and low-dimensional.

5. The nonlinear plasma response

5.1. Nonlinear effects

Linear and quasilinear approximations are useful for understanding qualitative dynamics of space plasmas, but not always optimized for reproducing the evolution or predicting it. More often than not, plasma systems are observed in a nonlinear state where instabilities are either saturated or quenched. The nonlinear states are described by wave-particle and wave-wave interactions where the higher-order terms in the interaction are too large to be neglected. Nonlinearity here refers to any violation of the superposition principle, either in the dynamics of an autonomous system or in the response of a driven system. The following examples illustrate the diversity of nonlinear dynamics and response.

Transpolar potential saturation. The solar wind-magnetosphere coupling is a complex interaction between a magnetized, cold plasma and a weakly magnetized, high-temperature plasma. As discussed in Section 4 the solar wind and IMF transfer energy and momentum to the high-latitude magnetosphere through reconnection. The IEF, parametrized by functions of the interplanetary plasma and magnetic field such as $V_{SW}B_{South}$, acts as an electric generator which produces the region-1 current and associated field-alined currents.

The electric potential across the polar cap, or transpolar potential, increases with IEF (Fig. 24). At low levels of the IEF (<3 mV/m) the response is linear, but for IEF>3 mV/m the transpolar potential, then ~100 keV, begins to saturate. The asymptotic potential is in the range 150-250 keV [Reiff and Luhmann, 1986]. The saturation of the potential has been explained as a negative feedback of the region-1 current on the rate of reconnection [Hill et al., 1976], but more recently as an effect of the ram pressure which limits the total current that can flow in the region-1 system [Siscoe et al., 2002]. In the latter case, the region-1 current takes the place of the magnetopause (Chapman-Ferraro) current in balancing the ram pressure via a **j**x**B** force.

Auroral electrojet response.

Studies of the magnetospheric activity in space and as reflected on geomagnetic activity on the ground show that there are two major modes of response. The first is a direct response to the solar wind energy input is manifested by the R1 current system and associated plasma convection. The second is a storage release, or loading-unloading, response [McPherron et al., 1973] which corresponds to the R2 current system. In the context of systems theory, a seminal study [Bargatze et al., 1985] examined the response of the westward current (using the AL electrojet index as a proxy for the actual current amplitude) to the solar wind input $V_{SW}B_{South}$. The response of the current showed clearly the different growth and decay timescales corresponding to the two types of magnetospheric response. The second, clearly nonlinear (storage-release) response has been identified with in-situ and remote-sensing techniques [Baker et al.,]. The key elements of the loading-unloading response have been reproduced in modern MHD simulations (Fig. 25). The results from the Bargatze et al. [1985] study motivated various parts of the nonlinear system theory and it is discussed in more detail at the end of Section 5.2.

Other salient examples of nonlinear response include: the response of the ring current dynamics and associated mid-latitude geomagnetic variations on the interplanetary inputs [Valdivia et al., 1999; Vassiliadis et al., 1999; O'Brien and McPherron, 2000]; the dependence of the magnetospheric electron flux on time variations of solar wind velocity and IMF [Paulikas and Blake, 1979; Blake et al., 1997, Li et al., 2001], etc.

5.2. Activity dependence and bifurcations.

Threshold models.

As the above example illustrate, nonlinearity can emerge is in response to an external driver. The dynamics is assumed to change when the driving goes from one level to the next, crossing an activity "threshold". Response functions or model coefficients are calculated for each level, using the FIR, AR, ARMA, and ODE models of Section 4.4. If the functions or coefficients change systematically, this is evidence that superposition of inputs does not hold. For a FIR model the change is of the form:

$$x(t) = x(t-T) + \int_0^T H_u^x(\tau;\bar{x})u(t-\tau)dt \tag{5.1}$$

where the impulse response function $H_u^x(\tau,\bar{x})$ is parametrized by a time average $\bar{x}$. Typically it is assumed that x(t-T)=0. The variation of $H_u^x(\tau,\bar{x})$ from one level to the next is compared to the uncertainty $|\delta H_i|$. The original FIR and Eq. (5.1) can also be compared via the validation statistics discussed in Sec. 4.4.

Example: Solar wind-westward electrojet coupling.

As discussed in 5.1, the high-latitude geomagnetic activity is mainly due to electrojet currents flowing horizontally in the ionosphere, and field-aligned currents that couple the magnetosphere to the high-latitude ionosphere. The driver for the electrojet currents is primarily the dayside magnetic reconnection, controlled by the IMF B_z component. Substorms and large-scale convection of various types (pseudobreakups, convection bays, SMC) are the main modes of response.

The interplanetary input $V_{SW}(t)B_s(t)$ is measured by an upstream solar wind monitor spacecraft. The amplitude of the electrojet activity is conventionally represented by the auroral electrojet index AL which measures the strongest Southward (geographic) deflection observed in a standard group of ground magnetometers. The index thus indirectly measures the intensity of the westward electrojet that responds to the solar wind input. In certain cases the network of magnetometers is suboptimal for recording the exact electrojet activity [Baumjohann, 1986]; however, it is instructive to study the index if it is treated carefully.

Bargatze et al. [1985] related the electrojet index output to the interplanetary field input by using a threshold MA model for a number of activity intervals of duration T = 2 days. Each interval was characterized by an integral occurrence probability of AL, and its activity level was parametrized by the median AL value (the AL corresponding to 50% of the occurrence probability), $med(AL)$. The activity-dependent response was thus:

$$AL(t) = \int_0^T H_{V_{SW}B_{South}}^{AL}(\tau;AL_{med})V_{SW}B_{South}(t-\tau)dt \tag{5.2}$$

Two of the impulse responses H=H(τ), for a low- and an intermediate-activity intervals are shown in Fig. 26. The time lag is measured relative to the solar wind impact at the magnetopause. The intermediate-level response $H_{V_{SW}B_{South}}^{AL}(\tau)$ has two peaks, indicating a strong response of the electrojet τ_1=20 minutes and then τ_2=60 minutes after solar wind impact. They correspond to the two components of the westward electrojet: the convection electrojet which is driven directly and the more irregular substorm electrojet in the midnight region, respectively.

A stack plot of impulse responses for all levels of activity is shown in Fig 27. The impulse response has the following qualitative structure

$$H_{V_{SW}B_{South}}^{AL}(\tau) = \begin{cases} Unimodal, \tau_1 \simeq 15\min, & AL_{med} < -90\,nT \\ Bimodal, \tau_1 \simeq 20\,and\,\tau_2 \simeq 60\min, & -90\,nT < AL_{med} < -40\,nT \\ Unimodal, \tau_1 \simeq 20\min, & -40\,nT < AL_{med} \end{cases} \tag{5.3}$$

The τ_1=20-min peak is the response of the Region-1 current system. At low AL levels only the direct response is observed: magnetic field line convection back to the dayside is completed fairly continuously. In the process, energy temporarily stored in the magnetotail, either as magnetic field energy in the lobes or thermal energy in the plasma sheet, is dissipated before it produces a substorm. At intermediate activity a second peak appears at τ_2=60 min. Convection is more impulsive and is accompanied by rapid dipolarization of the tail field and dissipation of the stored energy during substorms. The Region-2 current system contributes to the geomagnetic activity. For the strongest levels of driving, the convection occurs earlier, and the peak shifts to τ_1=15 min. The substorm electrojet disappears probably because a substorm current wedge does not have time to form properly. A second reason is the southward expansion of the auroral oval with activity, beyond the latitudes of the magnetometer network used to derive the AL index.

A comparison of the threshold FIR model (5.3) with a different type of model is instructive. An electric-circuit analogue in the form of a second-order linear ODE, was developed to reproduce the AL activity [Vassiliadis et al., 1993]. The analogue was driven with the $V_{SW}B_{South}$ time series data and optimized to fit the AL time series data from the Bargatze et al. [1985] dataset. The analogue's unimodal response function is analytically calculated and has the form of Eq. (4.24).

At low and high activity levels, the peak of the analogue response coincides with the τ_1 peak of the threshold model (5.3). At intermediate levels, the analogue's response cannot reproduce a second peak, but its single peak becomes broader with a width comparable to the combined duration of the τ_1 and τ_2 peaks of the threshold FIR model. In addition, the response time τ_1 is shortened for both models as activity level increases (cf. Eq. (5.3)). In terms of performance, the two models have nearly identical data-model correlations independently of activity level [Vassiliadis et al., 1995a].

Bifurcations.

While in threshold models the change of activity level determines the response, in a more fundamental approach to nonlinearity, a parameter may be used to change the equilibrium states of the system (e.g., from stable to unstable).

Many plasma modeling approaches assume the existence of a single, usually stable, equilibrium state which changes little, either in position or stability properties, as parameters are varied. Geospace plasmas, on the other hand, are notoriously volatile. Some of their large-scale variability can be explained in terms of a qualitative change in their equilibrium. A qualitative change in stability at a fixed parameter value is called a bifurcation; the parameter at which it occurs is called the threshold [Bergé et al., 1984]. A qualitative change here means a change in the type (stable or unstable), as well as the number, of equilibria.

The Hopf bifurcation is such a transition from a fixed point to an oscillation. The dynamics of the van der Pol oscillator of Eq. (3.14) is a function of the parameter ε: for $\varepsilon<0$ the asymptotic behavior of the system is the stable fixed point x(t)=0. As ε is increased, oscillatory activity in the form of a limit cycle appears at $\varepsilon = 0$.

Storage-release oscillators in space plasmas have a wide variety of dynamical regimes [Bernhardt, 1994]. The large-scale dynamics of the magnetosphere and its energy source (interplanetary medium) and sinks (ionosphere and plasma sheet) has been modeled in terms of large-scale electrodynamics with regional plasma physical corrections. The fundamental nonlinearity is in the magnetic lobes and plasma sheet and their capacitive and inductive properties [Klimas et al., 1992; 1994]. In the Faraday Loop model, lobe magnetic flux (and therefore stored magnetic energy) increases as magnetic field convects tailward after dayside reconnection. If the flux loading of Eq. (2.8) is slow, it is balanced by nightside reconnection and completion of the convection cycle. In addition, loading can occur following a thinning of the plasma sheet and in order to preserve pressure balance. When the lobe flux increases past a critical value, an instability develops and unloading begins. The loading and unloading rates are given by:

$$\begin{aligned} \frac{d\beta}{dt} &\propto E_{SW} \geq 0, \qquad && \Phi_{lobe}(t) < \Phi_c \\ \frac{d\beta}{dt} &= -\kappa \dot{\beta}\left(t_{\Phi=\Phi_c}\right) < 0, \quad && \Phi_{lobe}(t) > \Phi_c \end{aligned} \tag{5.4}$$

In the second equation of (5.4) the unloading rate is opposite to and proportional to the loading rate at the time when the lobe flux became critical, and much larger ($\kappa > 1$). This modeling assumption for the unloading rate is necessary in order for the model to be able to reproduce geomagnetic activity during periodic-substorm sequences as well as for its impulse response to resemble the impulse response function $H^{AL}_{V_{SW}B_{South}}(\tau)$ [Klimas et al., 1994; Freeman et al., 2004]. The state space of the model projected on the (E_{tail}, dE_{tail}/dt) plane is shown in Fig. 9.

In an alternative model, also tested on geomagnetic activity, the magnetotail's capacitive and inductive properties were modeled as the effects of mode conversion and absorption of low-frequency and the increase in plasma temperature and entropy which modifies the response to further waves [Goertz et al., 1991]. The dynamic equilibrium state of the model is a complex attractor, a projection of which in the (T,S) plane of the plasma-sheet state space is shown in Fig. 28. However, the model reproduces mainly the directly driven part of geomagnetic activity, i.e., the τ=20-min peak of $H^{AL}_{V_{SW}B_{South}}(\tau)$ [Goertz et al., 1993; McPherron and Rostoker].

The Faraday-Loop model was systematically developed to an electrodynamic model of the magnetotail and ionosphere with proper energy and current conservation [Horton and Doxas, 1996; 1998]. The coupled system is driven by the IEF $E_{ext}=V_{sw}B_{South}$ or a similar time-dependent input parameter. In abbreviated form, the Horton and Doxas model [1996; 1998] is given by:

$$\begin{aligned} L\frac{dI}{dt} &= -V + M\frac{dI_1}{dt} + l\cdot E_{ext}(t) \\ C\frac{dV}{dt} &= I - I_1 - I_{ps} - \Sigma V \\ \frac{3}{2}\frac{dP}{dt} &= \Sigma\frac{V^2}{\Omega} - u_0 K_{\|}^{1/2}\Theta(I - I_c)P \end{aligned} \tag{5.5}$$

where I(t) and V(t) are the cross-tail current and voltage; L, M, C, and l are the inductance, mutual inductance, plasma sheet capacitance, and solar wind coupling length scale; P is the plasma sheet ion pressure considered isotropic; $K_{\|}$ is the parallel kinetic energy due to precipitating electrons; and Ω is the magnetotail volume. Storage and release are represented by the term $\Theta(I-I_c)$ which is activated when the cross-tail current exceeds a critical value, corresponding to the threshold of an ion instability. The inductive-capacitive equations for the ionospheric current loop are analogous to the first two in (5.5).

Horton and coworkers [Horton et al., 2001; Smith et al., 2004] showed that model (5.5) goes through a complex series of bifurcations. As the activity level is raised, the system goes through increasingly complex equilibrium structures. Substorms are modeled as the transitions between these complex equilibria. At the highest driving amplitudes, the bifurcation tree becomes simpler in structure, and substorm sequences are eventually replaced by behavior that resembles convection bays, stormtime convection, or steady magnetospheric convection (SMC) events.

5.3. Chaos

In the previous sections, the plasma systems are stable under external perturbations, or changes in the state due to internal dynamics. Stability means that a small perturbation decays rather than grows. However, nonlinear plasmas often have equilibria that are unstable or marginally stable. Consider an autonomous (not driven) nonlinear system whose evolution is described by the nonlinear flow equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(\mathbf{x}) \tag{5.6}$$

Stability close to an equilibrium is examined by comparing the time evolution of two initially neighboring states $\mathbf{x}_1(0)$ and $\mathbf{x}_2(0)=\mathbf{x}_1(0)+\delta\mathbf{x}$. For small $|\Delta\mathbf{x}(0)| = |\delta\mathbf{x}|$, Eq. (5.6) can be linearized around $\mathbf{x}_1(0)$.

$$\frac{d\mathbf{x}}{dt} \simeq \mathbf{F}(\mathbf{x}_0) + \frac{\partial\mathbf{F}}{\partial\mathbf{x}}(\mathbf{x}-\mathbf{x}_0) = \mathbf{c} + \mathbf{A}\cdot\mathbf{x} \tag{5.7}$$

The ansatz is used $\mathbf{x} \propto \exp\left[\left(i\omega+\gamma\right)t\right]$. (The constant vector **c** is absorbed as a row in matrix **A**.) The eigenvalues of **A** are called Lyapunov exponents.[8] A positive Lyapunov exponent leads to an exponential increase between the $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$ trajectories along the corresponding eigenvector. Similarly, a negative exponent is associated with contraction along its eigendirection. Since contraction can be neglected, the net distance grows as

$$\left|\Delta\mathbf{x}\left(t\right)\right| \sim \left|\Delta\mathbf{x}\left(0\right)\right| e^{Kt} \tag{5.8}$$

where K, the Kolmogorov entropy, is the sum of the positive real exponents. If K is finite, a small difference in initial conditions increases exponentially.

Dynamical chaos is the regime in which essentially all trajectories in the state space are subject to the sensitivity of initial conditions (SIC), Eq. (5.8). Chaos strongly modifies the plasma response since it is a form of nonlinear feedback. Positive Lyapunov exponents are related to the system's instabilities.

The Kolmogorov entropy can be measured from experimental data. By initializing the system at two nearby positions in the reconstructed state space, and ensuring that other conditions (such as external driving) are practically identical, one can measure the SIC property.

In terms of forecasting, prediction of a chaotic system is limited beyond a timescale of the order 1/K. Increasing the prediction by one extra step, requires an exponential increase in the resolution of the initial conditions. Thus the prediction error grows exponentially:

$$e\left(t\right) \sim e\left(0\right) e^{Kt}$$

as indicated in Fig. 29, where e(0) is the initial error, which is typically the resolution of the measurements (compare with the linear growth of Eq. (4.31)). An example of measuring a positive Lyapunov exponent in a state space reconstructed from AL time series is shown in Fig. 30 (red curve).

Individual Lyapunov exponents correspond to "stretching and folding" of state space trajectories along the eigendirections, and can be measured by identifying directions of expansion and contraction [Bergé et al., 1984; Abarbanel et al., 1993]. The "stretch and fold" behavior leads to mixing of the plasma states. Thus chaos characterizes specific regions of the state space: depending on the dynamical parameters, there are regions of stable motion, for instance, islands around periodic equilibria, surrounded by chaotic areas and vice versa (as shown in Fig. 8 for motion close to a Harris sheet).

State space structure: attractors.

Since motion is bounded, divergence of solutions in one region of the state space must be followed by intervals of convergence. Eventually the system reaches an attractor to which it is henceforth confined. On the attractor, the time evolution **x**(t) may be periodic, quasiperiodic (characterized by several frequencies), or irregular (chaotic). Chaotic-dynamics attractors are much more complex than the fixed points and limit cycles of Section 4.

Negative Lyapunov exponents are the inverses of the times it takes to reach an attractor from specific directions. The exponents usually correspond to damping terms such as those in Eq. (3.2). If the dynamics are known, the exponents are calculated by linearizing the system equations and averaging over the trajectory segments; if it is not known, the damping timescales can be computed by data analysis, or by developing a state space model and measuring access times to the attractor. Fig. 30 shows an example of computing negative Lyapunov exponents for a state space reconstructed from the Bargatze et al. [1985] dataset. Rather than starting with a small error $e\left(0\right) \simeq 0$, the system starts from a large perturbation. The evolution of the error (black curves) indicates the effect of negative Lyapunov exponents, or of damping rates.

Chaotic system attractors typically have noninteger fractal dimension (in which case they are called "strange"). The dimension can be computed using the capacity dimension, d_2, or similar technique. Since the attractor is a subspace of the state space, d_2 is less than the dimension m of the surrounding state

[8] More precisely, Lyapunov exponents are defined in terms of n-th roots of the eigenvalues of the power matrix $\mathbf{A}^n$, corresponding to the n-th time step of evolution, with n tending to infinity.

space. Therefore, calculating the fractal dimension d_2, Eq. (3.18), of an attractor is a better estimate to df_{eff} than the dimension of the state space:

$$df_{eff} = \lceil d_2 \rceil \leq m \tag{5.9}$$

where $\lceil x \rceil$ denotes the smallest integer not exceeding x.

As discussed in Sec. 3.4, the dimension of magnetospheric activity was the focus of several studies in the early 1990s. Magnetospheric activity tends to returns to a quiescent state and the question was, whether the state could be described as an attractor. By reconstructing state spaces from time series of AE and other indices, it was possible to measure the fractal dimension of the attractor-like structure. The estimates of d_2 were in the range of 3-4 which was quite lower than the embedding dimension (10-15) [Vassiliadis et al., 1990; Roberts et al., 1991; Shan et al., 1991; Pavlos et al., 1999a, b]. However, the structure eventually was not classified as an attractor because the dimension estimators did not take into account the time-dependent interplanetary drivers.

Applications.

Examples of chaos in a Hamiltonian system include the charged particle motion in curved magnetic fields [Chen and Palmadesso, 1986; Büchner and Zelenyi, 1989]. Large-scale models of the magnetosphere dynamics can also be chaotic [Goertz et al., 1993; Klimas et al., 1992, 1994; Horton et al., 1996]. Nonlinear feedback mechanisms involve the plasma sheet resistivity as a function of the cross-tail current and the precipitating particle flux as a function of the cross-tail voltage. These internal instabilities should be distinguished from the role of the external drivers. For example, the magnetospheric response is widely different for similar initial conditions and solar wind and IMF of different amplitudes [Pulkkinen et al., 1998; Siscoe, 2002].

Chaos and turbulence.

Chaotic dynamics has several unique characteristics. It arises in nonlinear systems with very few degrees of freedom: three in a continuous-time system and two in a discrete-time system (a mapping). In addition, while df_{eff} is small, the power spectra of many chaotic-system time series were found to be broadband. This means that spectral techniques cannot measure the degrees of freedom of these chaotic systems correctly. Furthermore, it made chaos a candidate explanation for irregular, unpredictable activity of plasmas and fluids. The list included the global magnetospheric activity based on the results of the activity-dependent FIR model of Bargatze et al. [1985] and similar studies.

Chaotic dynamics has a complex relation to turbulence. In some systems, chaos arises because of turbulence (e.g., particle scattering in turbulent flow). In others, turbulence can lead to organization of the plasma system in modes which can then be destabilized leading to relaxation oscillations [Ortolani, 1989]; see Sec. 2.1). Yet in others, turbulence arises due to chaos (nonlinear interaction of a small number of large-scale modes in a fluid [Ruelle and Takens, 1971]. In the 1970s the chaos paradigm was put forward to explain the development of turbulence: a fluid or plasma system with a small number of effective degrees of freedom (<10) produces irregular, unpredictable patterns of activity in time. In a carefully chosen state space, these apparently irregular patterns often become recognizable as motion of the system state vector on an attractor. In that sense, chaos is an instability for a plasma system characterized by a few effective degrees of freedom. Although this type of behavior is locally unstable, it can be globally stable in the form of an attractor.

Some salient differences between dynamical chaos and turbulence are the following:

- Chaos produces a rapid (exponential) divergence of nearby trajectories, the "sensitivity to initial conditions" [Bergé et al., 1984]. Turbulence leads to a (nonlinear) diffusion in energy space.
- Chaos is identified in systems with small number of degrees of freedom, whereas turbulence is studied in many-degree-of-freedom fluids and plasmas. An intermediate category called spatiotemporal chaos, in which a spatially extended system displays dependence on initial conditions.
- Transitions from periodic to chaotic behavior ("routes to chaos") are well known. These transitions are different from transitions from laminar to turbulent behavior.

5.4. Nonlinear analysis tools

Local state space models.

In the threshold models described above, the activity levels are few and discrete. By increasing their number (i.e., decreasing the number of data per activity level, if the dataset is fixed) one can approximate a continuous representation of the nonlinear dynamics [Vassiliadis et al., 1995]. In this case the matrix that represents the local, linearized dynamics at the initial condition $\mathbf{x}(0)=\mathbf{x}_0$ becomes a continuous function of position in the state space:

$$\mathbf{A} = \mathbf{A}(\mathbf{x}_0)$$

The matrix **A** is computed from (5.7) using state vectors reconstructed from experimental time series. The vectors are selected to be similar to $\mathbf{x}_0$ and therefore distributed around it in the state space, and are called its (nearest) neighbors.

The state space model is thus local, or valid for a small neighborhood of the state space. In contrast, all linear models of the previous section were global. Geometrical methods that can be applied make these models attractive. For example the degree of similarity between states $\mathbf{x}(t_i)$ and $\mathbf{x}(t_j)$ is represented by the distance:

$$d_{ij} = \left|\mathbf{x}(t_i) - \mathbf{x}(t_j)\right|$$

One way to select neighbors of $\mathbf{x}_0$ is by distance, i.e. the neighbors are those points

$$\left\{\mathbf{x}(t_i) \mid d_{i0} < r\right\} \tag{5.10}$$

Properties of the model are expressed in terms of the distribution of neighbors. For example, the number of nearest neighbors, NN, is the sum:

$$NN(\mathbf{x}_0; r) = \sum_i^N \Theta(d_{i0} < r) \tag{5.11}$$

and the average state is the ensemble average

$$\left\langle \mathbf{x}(t_i) \right\rangle_{\{\mathbf{x}(t_i)|d_{i0}<r\}}$$

The local model has the form of Eq. (5.7), but parameters **c** and **A** are computed only from the nearest neighbors rather than the entire time series data:

$$\frac{d\mathbf{x}}{dt} = \mathbf{c}(\mathbf{x}_0; NN) + \mathbf{A}(\mathbf{x}_0; NN) \cdot \mathbf{x} \tag{5.12}$$

Stopping at the zeroth-order term **c** gives a local-constant model while keeping **c** as well as **A** produces a local-linear model. (Fig. 31). Local-linear models are more accurate than linear models as seen in studies of the electrojet indices as indicated also in Fig. 23 [Vassiliadis et al., 1995a,b]. However, their response functions are more noisy that those of the simpler FIR and ARMA models, and are harder to interpret physically.

The specification criterion (5.10) gives significant flexibility in modeling. The magnitude $\left|\mathbf{x}(t)\right|$ represents the activity level while r represents the state space resolution of interest. If the plasma response is independent of r, and thus of the activity level, it is linear. Also if r is of the size of the radius of the state space, ensemble averages like (5.11) tend towards the time-average $\overline{x}$. The nearest-neighbor specification (5.10) can be used to make more precise the traditionalsuperposed-epoch analysis

$$\left\{\mathbf{x}(t_i) \mid i : Criteria\right\}$$

where "Criteria" stands for conditions on the state vectors $\mathbf{x}$(t).

Conditional averaging. Nearest-neighbor techniques belong to techniques of conditional averaging. In the model (5.12), the condition is $d_{i0}<r$. Conditional averaging can be used to extract modes of response from noisy experimental data such as those of two-dimensional electrostatic plasma turbulence

[Pécseli and Trulsen, 1991; Nielsen et al., 1996]. In addition to the usefulness of geometrical methods, if the model order is low, one can use visualization techniques for the state space as well as the dynamics.

Neural networks.

In contrast to local-linear models, neural networks are global-nonlinear models and can be considered as generalizations of the FIR and in some cases the ARMA-type models. However, there are two major differences between neural networks and all other models discussed so far, in addition to the nonlinearity. The first is the layered structure, or "architecture", and the second the iterative calculation of model coefficients. ODEs and ARMAs solve the system equations directly, while neural networks reach the optimal model coefficients iteratively.

Introduced in the 1960s and conceptually based on the communication among brain neurons, neural, or "connectionist," networks are multi-layer models with adjustable coupling between elements of different layers [Hertz et al., 1991]. As the model is trained on combinations of input and output variables, the connections are adjusted so that the network reproduces the system output as closely as possible. An important result of neural network theory is that a two-layer network can reproduce any smooth input-to-output function given a sufficient amount of low-noise data.

Here, too, the modeling approach relies on a state space representation. Networks are used to classify and model high-dimensional data such as 2-dimensional images. For time-series variables the network approach uses state vectors such as

$$\mathbf{x}(t)=\left[x(t),x(t-\tau),x(t-2\tau),\ldots,x\left(t-(m-1)\tau\right)\right] \tag{5.13}$$

similar to the state vector of the threshold FIR model of Section 5.2. A two-layer perceptron relates an input **u**(t) to the scalar state, or output, x(t) via an intermediate, "hidden" layer $h_i(t)$:

$$\begin{aligned} x(t) &= g\left(\mathbf{h}(t)\right)=g\left(\sum_{k=1}^{m} w_{ik}^{(1)}{}_{ik}h_i-\theta_i^{(1)}\right) \\ h_i(t) &= g\left(\mathbf{u}(t)\right)=g\left(\sum_{k=1}^{m} w_{ik}^{(0)}u(t-k\tau)-\theta_i^{(0)}\right) \end{aligned} \tag{5.14}$$

Layers are indicated by a superscript (0 or 1). The nonlinear activation function g of a network node is often selected to be a sigmoid function such as the hyperbolic tangent. If g is the identity function, $g(h_i)=h_i$, the network becomes the FIR model (Eq. 4.22). The input to the i-th element must be greater than a bias θ_i for the neuron to respond.

Both the weights w_{ik} and the biases θ_i at each node (5.14) are adjusted iteratively as functions of the rms prediction error (4.16). The iterative solution of neural networks produces more accurate models of nonlinear systems compared to direct inversions used in ARMA or ODE models [Hertz et al., 1991].

In recurrent networks, the feedforward architecture of neural networks such as that of Eq. (5.14) is generalized to include feedback from earlier outputs such as $x(t-1)$. Elman recurrent networks can lead to significant improvement in data-model correlation [e.g., Wu et al., 1996].

Advantages and limitations.

Neural networks are nonlinear adaptive models which generalize FIR and ARMA models. Their structural flexibility leads to higher accuracy than their linear counterparts.

The main limitation of networks is the difficulty in physically interpreting the weights and biases, especially if the number of layers exceeds 2. However, several approaches have been proposed to extract physical information from networks. First, it is possible to examine their output under controlled, or parametrized, inputs [O'Brien and McPherron, 2000]. A special case of this approach is to drive the network with impulses of different amplitudes to obtain an impulse response [Weigel et al., 1999]. The impulse response peak time and amplitudes and their dependence on the input amplitude are used to identify the physical processes that the network represents. Second, for recurrent networks, context nodes (those that couple to earlier outputs such as x(t-1)) correspond to lags in (5.13) and can be interpreted as effective degrees of freedom. In a study of the AL index Gleisner et al. [1997] showed that the number of context nodes controls the network's prediction accuracy.

Applications.

A growing number of studies have used neural networks to classify and model geospace physics data [Joselyn et al., 1994; Sandahl and Jonsson, 1998]. There is considerable potential for neural networks and other methods of data mining since the volume of space physics databases increases faster than can be analyzed.

The first geospace study to use networks, by Koons and Gorney [1991], showed that feedforward neural networks are significantly more accurate than linear MA models in reproducing and forecasting relativistic electron fluxes at geosynchronous orbit from solar wind input parameters.

Most of the applications are in modeling geomagnetic indices from solar wind parameters. A series of studies has focused on the time variation of the D_{st} index as a proxy for the dynamics of the ring current and related current systems [Wu and Lundstedt, 1996, 1997]. Jankovicova et al. [2002] optimized the solar wind input to their D_{st} neural network by using a principal component basis. High-latitude geomagnetic activity has also been studied with neural networks including the response of geomagnetic indices [Gleisner et al., 1997; Takalo et al., 1999; Weigel et al., 1999], the regional magnetic field [Sutcliffe, 2000], and substorm occurrence [Sutcliffe, 1997].

6. **Summary, conclusions, and outlook**

This review of systems theory for geospace plasmas presented the powerful systems approach and discussed some of the ways it has been, and can further be, combined with traditional plasma physics theory. Since this is a tutorial text, several special topics had to be left out or simply referenced. In particular, the modeling techniques for spatially extended systems, a rapidly growing discipline, following the advent of multi-satellite missions and the development of ground arrays around the world, is deferred to a future paper.

Systems applications has provided a foundation for space weather and space engineering. In addition, two different developments take place in the numerical arena, where the increase in computational power is such that planetary-scale codes can be used to represent events originating at the Sun and reaching Earth's lower atmosphere [De Zeeuw]. First, these simulations have started assimilating observational data, and eventually will do so in real time. Such efforts have been already initiated in ionospheric and radiation belt models [Ridley; Schunk; Rigler et al.]. Judging from the developments in meteorology and oceanography, data assimilation will be pivotal in the next generation of systems applications. Second, in the near future, as simulations become increasingly realistic, the volumes of on-line high-quality data will rapidly surpass those of observations. In several years simulations will be sufficiently realistic that their outputs can be used for data analysis studies in their own right. This meta-modeling will rely on nonlinear systems theory rather than microscopic physics in order to expand our understanding of the complex dynamics in simulations.

Acknowledgments. This tutorial review is based on lectures given at the School of Analysis Techniques for Space Plasma Data held in La Londe, Les Maures, France, on October 8-13, 2001. I am thankful to the organizers, Thierry Dudok de Wit and Joachim Vogt, for their initiative and hard work in implementing the school. At the time of writing, lectures and other presentations are archived at http://lpce.cnrs-orleans.fr/~ddwit/lalonde/index.html. I also thank A. Klimas, D. Baker, and R. Weigel for many discussions over the last decade on nonlinear modeling of space plasmas. Many thanks are due to R.L. McPherron, D. Papadopoulos, S. Sharma, R. Ghosh, V. Uritsky, A. Pulkkinen, and H. Pécseli for discussions. Finally I am thankful to P. Cargill and the RoG editors for giving me the space to create this synthesis.

List of Symbols and Notation:

Roman symbols

a_i	Coefficient of an AR or ARMA model
AL, AU, AE	Auroral electrojet (AE) geomagnetic indices
b_i	Coefficient of an MA or ARMA model
B	Magnetic field
B_{North}, B_{South}	Northward (positive), Southward (negative) components of the IMF
B_T	Transverse component of the IMF, i.e. in the y-z plane.
$C_X(\tau)$	Autocorrelation function of X(t) as a function of delay time τ.
$C_{X,Y}(\tau)$	Cross-correlation function of X(t) and Y(t) as a function of delay time τ.
	$C_X(0)$ is the linear correlation coefficient.
$\mathbf{C_x}$	Covariance matrix of **x** (which can be multivariate).
CI	Correlation integral.
D_{st}	Geomagnetic index D_{st}
d	Fractal dimension
d_0	Capacity dimension
d_2	Correlation dimension
d_q	the q-th dimension in the multifractal spectrum
E	Electric field
E	a) Energy
	b) Electric field magnitude
e_{rms}	Root-mean-square error
f	a) Frequency
	b) Phase space distribution
$H_X^Y(\tau)$	FIR filter: Impulse response function of Y(t) to X(t) as a function of delay time τ
J	Particle flux
j	Current density
K	Kolmogorov entropy
K_p	K_p geomagnetic index
L	L shell
n	number density
P	Pressure tensor
P(x)	Probability density function of x
S	entropy
T	a) Time; duration of an interval
	b) Temperature
v	Velocity
V_{SW}	Solar wind bulk speed
W	Singular value matrix
w_i	i-th singular value

Greek Symbols

α	Pitch angle
Θ	Heaviside, or step, function
λ_i	i-th Lyapunov exponent
σ_x	Standard deviation of x
τ	Lag time
τ_{loss}	Loss time scale
Φ	magnetic flux through a surface
ω	Circular frequency, $2\pi f$

Notation

x,y,z,… Scalar variables
x,**y**,**z**,… Vector variables

A,**B**,**C**,…	Matrices
$\mathbf{A}^{\dagger}$	Transpose of matrix **A**
$\lceil x \rceil$	Smallest integer not exceeding x
$\overline{x}$	Time-average of x(t)
$\langle x \rangle$	Ensemble average of x
$\tilde{x}(\omega)$	Fourier, Laplace, or z transform of x(t)
$\hat{x}$	Prediction for x

References

Review articles:

Abarbanel, H.D.I., R. Brown, J.J. Sidorowich, L.S. Tsimring, The analysis of chaotic data in physical systems, Rev. Mod. Phys. 65, 4, 1331-1392, 1993.

Consolini, G., and T.S. Chang, Complexity, magnetic field topology, criticality, and metastability in magnetotail dynamics, J. Atmos. Sol.-Terr. Phys. 64, 5-6, 541-549, 2002.

Horton, W., Chaos and structures in the magnetosphere, Phys. Rep.-Phys. Lett. 283, 1-4, 265-302, 1997.

Klimas, A.J., D. Vassiliadis, D.N. Baker, D.A. Roberts, The organized nonlinear dynamics of the magnetosphere, J. Geophys. Res. 101, A6, 13,089-13,113, 1996.

Sharma, A.S., Assessing the magnetosphere's nonlinear behavior – its dimension is low, its predictability is high, Rev. Geophys. 33, 645-650, Part 1, Suppl. S, 1995.

Song, P., and C.T. Russell, Time series data analyses in space physics, Space Sci. Rev. 87, 387-463, 1999.

Vassiliadis, D., A.J. Klimas, J.A. Valdivia, D.N. Baker, The nonlinear dynamics of space weather, Space Weather: Physics and Applications, Adv. Space Res. 26, 1, 197-207, 2000.

Regular articles:

Akaike, H., Canonical correlation analysis of time series and the use of an information criterion, in: *System identification: advances and case studies*, R. Mehra and D. Lainiotis (eds.), Academic Press, New York, NY, 1976.

Anderson, B.J., M.J. Engebretson, S.P. Bounds, L.J. Zanetti, T.A. Potemra, A statistical study of Pc-3-5 pulsations observed by the AMPTE CCE magnetic fields experiment. 1. Occurrence distributions, J. Geophys. Res. 95, A7, 10,945-10,523, 1990.

Arfken, G.B., and H.-J. Weber, *Mathematical Methods for Physicists*, 4th edition, Academic Press, San Diego, CA, 1995.

Baker, D.N., Statistical analyses in the study of solar wind-magnetosphere coupling, in: *Solar Wind-Magnetosphere Coupling*, Y. Kamide and J.A. Slavin (editors), Terra Scientific, Tokyo, 1986.

Baker, D. N., R. L. McPherron, T. E. Cayton, R. W. Klebesadel, Linear prediction filter analysis of relativistic electron properties at 6.6 RE, J. Geophys. Res. 95, A9, 15,133-15,140, 1990.

Baker, D.N., A.J. Klimas, R.L.McPherron, J. Büchner, The evolution from weak to strong geomagnetic activity – an interpretation in terms of deterministic chaos, Geophys. Res. Lett. 17, 1, 41-44, 1990.

Baker, D.N., S.G. Kanekal, A.J. Klimas, D. Vassiliadis, T.I. Pulkkinen, Collective phenomena in the inner magnetosphere, Phys. Plasmas 6, 11, 4195-4199, 1999.

Balikhin, M.A., O.M. Boaghe, S.A. Billings, H.S.K. Alleyne, Terrestrial magnetosphere as a nonlinear resonator, Geophys. Res. Lett. 28, 6, 1123-1126, 2001a.

Balikhin, M.A., I. Bates, S. Walker, Identification of linear and nonlinear processes in space plasma turbulence data, in: Alfvenic structures: From the Sun to the Magnetosphere, Adv. Space Res. 28, 5, 787-800, 2001b.

Bargatze, L.F., D.N. Baker, R.L. McPherron, E.W. Hones, Magnetospheric impulse response for many levels of geomagnetic activity, J. Geophys. Res. 90, A7, 6387-6394, 1985.

Baumjohann, W., Advantages and limitations of magnetospheric indices, in: *Solar Wind-Magnetosphere Coupling*, Y. Kamide and J.A. Slavin (eds.), Terra Scientific, Tokyo, 1986.

Behnke, R., et al., *The National Space Weather Program: The Implementation Plan*, Office of the Federal Coordinator for Meteorological Services and Supporting Research, FCM-P31-1997, Washington, DC, 1997.

Bendat, J.S., and A.G. Piersol, *Random data: analysis and measurement procedures*, 2nd edition, Wiley, New York, NY, 1986.

Bergé, P., Y. Pomeau, C. Vidal, *Order within Chaos: Towards a Deterministic Approach to Turbulence*, Wiley, New York, 1984.

Bevington, P.E., and D.K. Robinson, *Data reduction and error analysis for the physical sciences*, 2nd ed., McGraw-Hill, New York, NY, 1992.

Birdsall, C.K., and A.B. Langdon, *Plasma Physics Via Computer Simulation*, Institute of Physics Publishing, 1991.

Biskamp, D., *Nonlinear Magnetohydrodynamics*, Cambridge University Press, Cambridge, UK, 1993.

Blake, J.B., D.N. Baker, N. Turner, K.W. Ogilvie, R.P. Lepping, Correlation of changes in the outer-zone relativistic-electron population with upstream solar wind and magnetic field measurements, Geophys. Rev. Lett. 24, 8, 927-929, 1997.

Boberg, F., P. Wintoft, H. Lundstendt, Real time K_p predictions from solar wind data using neural networks, Phys. Chem. Earth C, 25, 4, 275-280, 2000.

Book, D.L., The NRL plasma formulary, , ***2000.

Bothmer, V., and R. Schwenn, The structure and origin of magnetic clouds in the solar wind, Ann. Geophys. 16, 1, 1-24, 1998.

Büchner, J., and L.M. Zelenyi, Regular and chaotic charged-particle motion in magnetotail-like field reversals. 1. Basic theory of trapped motion, J. Geophys. Res. 94, A9, 11,821-11,842, 1989.

Burkhart, G.R., J.F. Drake, P.B. Dusenbery, T.W. Speiser, Particle model for magnetotail neutral current sheet equilibria, J. Geophysical Research 97, A9, 13,799-13,815, 1992.

Burlaga, L.F., Multifractal structure of the magnetic field and plasma in recurrent streams at 1 AU, J. Geophys. Res. 97, A4, 4283-4293, 1992.

Burlaga, L.F., Lognormal and multifractal distributions of the heliospheric magnetic field, J. Geophys. Res. 106, A8, 15,917-15,927, 2001.

Burton, R.K., L.R. McPherron, C.T. Russell, Empirical relationship between interplanetary conditions and Dst, J. Geophys. Res. 80, 31, 4204-4214, 1975.

Casdagli, M., and S. Eubank, *Nonlinear Modeling and Forecasting*, Addison-Wesley, Redwood City, CA, 1992.

Chang, T., Low-dimensional behavior and symmetry-breaking of stochastic systems near criticality – can these effects be observed in space and in the laboratory?, IEEE Trans. Plasma Sci. 20, 6, 691-694, 1992.

Chapman, S.C., N. W. Watkins, R. Dendy, P. Helander and G. Rowlands, A simple avalanche model as an analogue for magnetospheric activity, Geophys. Res. Lett., 25, 2397, 1998.

Chen, F.F., *Introduction to Plasma Physics and Controlled Fusion*, 2nd edition, Plenum, New York, 1984.

Chen, J., and P.J. Palmadesso, Chaos and nonlinear dynamics of single-particle orbits in a magnetotaillike magnetic-field, J. Geophys. Res.-Space Phys. 91, A2, 1499-1508, 1986.

Chi, P.J., and C.T. Russell, An interpretation of the cross-phase spectrum of geomagnetic pulsations by the field line resonance theory, Geophys. Res. Lett. 25, 24, 4445-4448, 1998.

Chui, C.K., *An Introduction to Wavelets (Wavelet Analysis and its Applications, Vol. 1),* Academic Press, Boston, MA, 1992.

Clauer, C.R., The technique of linear prediction filters applied to studies of solar wind-magnetosphere coupling, in: *Solar Wind-Magnetosphere Coupling*, Y. Kamide and J.A. Slavin (editors), Terra Scientific, Tokyo, 1986.

Consolini, G., M.F. Marcucci, and H. Kandidi, Multifractal structure of auroral electrojet index data, Phys. Rev. Lett. 76, 21, 4082-4085, 1996.

Degtjarev, V.I., G.V. Popov, S.E. Chudnenko, Effect of the magnetic activity prehistory on the electron flux levels in geostationary orbit, Geomag. Aeron. 42, 3, 299-304, 2002.

De Wit T.D., V.V. Krasnoselkikh, Wavelet bicoherence analysis of strong plasma turbulence at the Earth's quasiparallel bow shock, Phys. Plasmas 2, 11, 4307-4311, 1995.

De Wit, T.D., V.V. Krasnoselkikh, T. Dunlop, H. Luhr, Identifying nonlinear wave interactions in plasmas using two-point measurements: A case study of Short Large Amplitude magnetic Structures (SLAMs), J. Geophys. Res. 104, A8, 17,079-17,090, 1999.

Doxas, I., W. Horton, and R.S. Weigel, Using particle simulations for parameter tuning of dynamical models of the magnetotail, J. Atmos. Sol.-Terr. Physics 64, 5-6, 2002.

Dunlop, M.W., T.I. Woodward, and C.J. Farrugia, Minimum Variance Analysis: Cluster themes, in: *Proceedings of the Cluster Workshop on Data Analysis Tools, Braunschweig, Germany, 28-30 September 1994*, K.-H Glassmeier and U. Motschmann (editors), SP-371, ESA, Noordwijk, Holland, 1995.

Engebretson, M., K.-H. Glassmeier, M. Stellmacher, W.J. Hughes, H. Lühr, The dependence of high-latitude Pc5 wave power on solar wind velocity and on the phase of high-speed solar wind streams, J. Geophys. Res. 103, A11, 26,271-26,283, 1998.

Farrugia, C.J., V.A. Osherovich, and L.F. Burlaga, Magnetic-flux rope versus the spheromak as models for interplanetary magnetic clouds, J. Geophys. Res. 100, A7, 12,293-12,306, 1995.

Fraser, A.M., and H. L. Swinney, Independent coordinates for strange attractors from mutual information, Phys. Rev. A 33, 2, 1134-1140, 1986.

Ghosh, S., and W.H. Matthaeus, Relaxation processes in a turbulent compressible magnetofluid, Phys. Fluids B2, 7, 1520-1534, 1990.

Glassmeier, K.-H., U. Motschmann, R. Vonstein, Mode recognition of MHD wave-fields at incomplete dispersion measurements, Ann. Geophys. 13, 1, 76-83, 1995.

Gleisner, H., H. Lundstedt, P. Wintoft, Predicting geomagnetic storms from solar-wind data using time-delay neural networks, Ann. Geophys. 14, 7, 679-686, 1996.

Gleisner, H., H. Lundstedt, Response of the auroral electrojets to the solar wind modeled with neural networks, J. Geophys. Res. 102, A7, 14,269-14,278, 1997.

Goertz, C.K., R.A. Smith, L.-H. Shan, Chaos in the plasma sheet, Geophys. Res. Lett. 18, 8, 1639-1642, 1991.

Goertz, C.K., L.-H. Shan, R.A. Smith, Prediction of geomagnetic activity, J. Geophys. Res. 98, A5, 7673-7684, 1993.

Goldstein, H., *Classical Mechanics*, Addison-Wesley, Reading, MA, 1980.

Grassberger, P., and I. Procaccia, Measuring the strangeness of strange attractors, Phys. D 9, 1-2, 189-208, 1983.

Hasegawa, A., Self-organization processes in continuous media, Adv. Phys. 34, 1, 1-42, 1985.

Hernandez, J., W. Horton, and T. Tajima, Low-frequency mobility response functions for the central plasma sheet with application to tearing modes, J. Geophys. Res. 98, A4, 5893-5906, 1993.

Hertz, J., A. Krogh, R.G. Palmer, *Introduction to the Theory of Neural Computation*, Cambridge University Press, Cambridge, UK, 1991.

Higuchi, T. Relationship between the fractal dimension and the power-law index for a time series – a numerical investigation, Physica D 46, 2, 245-264, 1990.

Horton, W., and I. Doxas, A low-dimension energy-conserving state space model for substorm dynamics, J. Geophys. Res. 101, A12, 27,223-27,237, 1996.

Horton, W., and I. Doxas, A low-dimensional dynamical model for the solar wind driven geotail-ionosphere system, J. Geophys. Res. 103, A3, 4561-4572, 1998.

Horton, W., J.P. Smith, R. Weigel, C. Crabtree, I. Doxas, B. Goode, J. Cary, The solar-wind driven magnetosphere-ionosphere as a complex dynamical system, Phys. Plasmas 6, 11, 4178-4184, 1999.

Horton, W., R.S. Weigel, and J.C. Sprott, Chaos and the limits of predictability for the solar-wind-driven magnetosphere-ionosphere system, Phys. Plasmas 8, 6, 2946-2952, 2001.

Horton, W., R.S. Weigel, D. Vassiliadis, and I. Doxas, Substorm classification with the WINDMI model, Nonlin. Proc. Geophys. 10, 4-5, 363-371, 2003.

Howell, K.C., B.T. Barden, M.W. Lo, Application of dynamical systems theory to trajectory design for a libration point mission, J. Astron. Sci. 45, 2, 161-178, 1997.

Ichimaru, S., Basic Principles of Plasma Physics: A Statistical Approach, Addison-Wesley, 1973.

Iyemori, T., H. Maeda, and T. Kamei, Impulse response of geomagnetic indices to interplanetary magnetic field, J. Geomag. Geoelectr. 31, 1, 1-9, 1979.

Jackson, J.E., *A user's guide to principal components*, Wiley, New York, NY, 1991.

Jankovicova, D., P. Dolinsky, F. Valach, Z. Voros, Neural network-based nonlinear prediction of magnetic storms, J. Atm. Sol.-Terr. Phys. 64, 5-6, 651-656, 2002.

Joselyn, J., H. Lundstedt, J. Trollinger, *Proceedings of the International Workshop on Artificial Intelligence Applications in Solar-Terrestrial Physics, Lund, Sweden, 22-24 September 1993*, NOAA, Boulder, CO, 1994.

Joyce, D., and D. Montgomery, J. Plasm. Phys. 10, 107, 1973.

Kanekal, S.G., D.N. Baker, J.B. Blake, B. Klecker, R.A. Mewaldt, G.M. Mason, Magnetospheric response to magnetic cloud (coronal mass ejection) events: Relativistic electron observations from SAMPEX and Polar, J. Geophys. Res. 104, A11, 24,885-24,894, 1999.

Kanekal S.D., D.N. Baker, J.B. Blake, Multisatellite measurements of relativistic electrons: global coherence, J. Geophys. Res. 106, A12, 29,721-29,732, 2001.

Kennel, M.B., R. Brown, H.D.I. Abarbanel, Determining embedding dimension for phase-space reconstruction using a geometrical construction, Phys. Rev. A 45, 6, 3403-3411, 1992.

Klimas, A.J., D.N. Baker, and D.A. Roberts, Linear prediction filters for linear and nonlinear modeled geomagnetic activity, Geophys. Res. Lett. 18, 8, 1635-1638, 1991.

Klimas, A.J., D.N. Baker, D.A. Roberts, D.H. Fairfield, J. Buchner, A nonlinear dynamics analog model of geomagnetic activity, J. Geophys. Res. 97, A8, 12,253-12,266, 1992.

Klimas, A.J., D.N. Baker, D. Vassiliadis, and D.A. Roberts, Substorm recurrence during steady and variable solar-wind driving – evidence for a normal mode in the unloading dynamics of the magnetosphere, J. Geophys. Res. 99, A8, 14,855-14,861, 1994.

Klimas, A.J., D. Vassiliadis, D.N. Baker, Data-derived analogues of the magnetospheric dynamics, J. Geophys. Res. 102, A12, 26,993-27,009, 1997.

Klimas, A.J., D. Vassiliadis, D.N. Baker, Dst index prediction using data-derived analogues of the magnetospheric dynamics, J. Geophys. Res. 103, A9, 20,435-20,447, 1998.

Klimas, A.J., D. Vassiliadis, D.N. Baker, J.A. Valdivia, Data-derived analogues of the solar wind-magnetosphere interaction, Phys. Chem. Earth C, 24, 1-3, 37-44, 1999.

Koons, H.C., and D.J. Gorney, A neural network model of the relativistic electron flux at geosynchronous orbit, J. Geophys. Res. 96, 5549-5556, 1991.

Kraichnan, R.H., Inertial ranges in 2-dimensional turbulence, Phys. Fluids 10, 7, 1417, 1967.

Kudo, K., Fractal analysis of electrical trees, IEEE Trans. Diel. Electr. Ins., 5, 5, 713-727, 1998.

Lefeuvre, F., V. Kravtchenkko-Berenoj, and D. Lagoutte, Introduction to the analysis of bi-linear and tri-linear processes in space plasmas, in: *Proceedings of the Cluster Workshop on Data Analysis Tools, Braunschweig, Germany, 28-30 September 1994*, K.-H. Glassmeier and U. Motschmann (editors), SP-371, ESA, Noordwijk, Holland, 1995.

Lin, R.P., R.A. Schwartz, S.R. Kane, R.M. Pelling, K.C. Hurley, Solar hard X-ray microflares, Ap. J. 285, 421, 1984.

Lo, M.W., B.G. Williams, W.E. Bollman, D.S. Han, Y.S. Hahn, J.L. Bell, E.A. Hirst, R.A. Corwin, P.E. Hong, K.C. Howell, B. Barden, R. Wilson, Genesis mission design, J. Astron. Sci. 49, 1, 169-184, 2001.

Lopez, R.E., M. Wiltberger, S. Hernandez, J.G. Lyon, Solar wind density control of energy transfer to the magnetosphere, Geophys. Res. Lett. 31, 8, Art. No. L08804, 2004.

Lui, A.T.Y., R.E. Lopez, S.M. Krimigis, R.W. McEntire, L.J. Zanetti, T.A. Potemra, A case study of magnetotail current sheet disruption and diversion, Geophys. Res. Lett. 15, 7, 721-724, 1988.

Lui, A.T.Y., P.H. Yoon, C.L. Chang, Quasi-linear analysis of ion Weibel instability in the Earth's neutral sheet, J. Geophys. Res. 98, A1, 153-163, 1993.

Lui, A.T.Y., S.C. Chapman, K. Liou, P.T. Newell, C.-I. Meng, M. Brittnacher, G.D. Parks, Is the dynamic magnetosphere an avalanching system?, Geophys. Res. Lett. 27, 7, 911-914, 2000.

Lui, A.T.Y., Multifractal and intermittent nature of substorm-associated magnetic turbulence in the magnetotail, J. Geophys. Res. 63, 13, 1379-1385, 2001.

Lyons, L.R., and D.J. Williams, *Quantitative aspects of magnetospheric physics*, Dordrecht, Boston, MA, 1984.

Mandelbrot, B.B., *The Fractal Geometry of Nature*, Freeman, 1983.

Matthaeus, W.H., and M.L. Goldstein, Measurement of the rugged invariants of magneto-hydrodynamic turbulence in the solar wind, J. Geophys. Res. 87, NA8, 6011-6028, 1982.

Matthaeus, W.H., and D. Montgomery, Dynamic alignment and selective decay in MHD, 285-291, in: *Statistical Physics and Chaos in Fusion Plasmas*, C.W. Horton, Jr., and L.E. Reichl, Wiley, New York, 1984.

McPherron, R.L., C.T. Russell, and M.P. Aubry, Satellite studies of magnetospheric substorms on August 15, 1968. Phenomenological model for substorms, J. Geophys. Res. 78 (16): 3131-3149 1973.

Milovanov, A.V., L.M. Zelenyi, G. Zimbardo, P. Veltri, Self-organized branching of magnetotail current systems near the percolation threshold, J. Geophys. Res. 106, A4, 6291-6307, 2001.

Motschmann, U., and K.-H. Glassmeier, Mode recognition of MHD wave fields, in: Proceedings of the Cluster Workshop on Data Analysis Tools, Braunschweig, Germany, 28-30 September 1994, K.-H. Glassmeier and U. Motschmann (eds.), ESA-SP-371, ESA, Noordwijk, 1995.

Motschmann, U., T.I. Woodward, K.-H. Glassmeier, D.J. Southwood, J.L. Pincon, Wavelength and direction filtering by magnetic measurements at satellite arrays: Generalized minimum variance analysis, J. Geophys. Res. 101, A3, 4961-4965, 1996.

Moore, C., Braids in classical dynamics, Phys. Rev. Lett. 70, 24, 3675-3679, 1993.

Nagai, T., "Space weather forecast:" prediction of relativistic electron intensity at synchronous orbit, Geophys. Res. Lett. 15, 425-428, 1988.

Newman, D.L., M.V. Goldman, R.E. Ergun, A. Mangeney, Formation of double layers and electron holes in a current-driven space plasma, Phys. Rev. Lett. 87, 25, 2001.

Nicolis, G., and I. Prigogine, *Exploring complexity: an introduction*, Freeman, New York, NY, 1989.

Nielsen, A.H., H.L. Pécseli, J.J. Rasmussen, Turbulent transport in low-β plasmas, Phys. Plasmas 3, 5, 1530-1544, 1996.

Niemeyer, L., L. Pietronero, H.J. Wiesmann, Fractal dimension of dielectric-breakdown, Phys. Rev. Lett. 52, 12, 1033-1036, 1984.

O'Brien, T.P., and R.L. McPherron, An empirical phase space analysis of ring current dynamics: Solar wind control of injection and decay, J. Geophys. Res. 105, A4, 7707-7719, 2000.

O'Brien, T.P., and R.L. McPherron, A neural network model relating Delta H at a single station to Dst, in: Space Weather: Physics and Applications, Adv. Space Res. 26, 1, 209-212, 2000.

Ohtani, S., T. Higuchi, A.T.Y. Lui, K. Takahashi, Magnetic fluctuations associated with tail current disruption – fractal analysis, J. Geophys. Res. 100, A10, 19,135-19,145, 1995.

Ohtani, S., K. Takahashi, T. Higuchi, A.T.Y. Lui, H.E. Spence, J.F. Fennell, AMPTE/CCE-SCATHA simultaneous observations of substorm-associated magnetic fluctuations, J. Geophys. Res. 103, A3, 4671-4682, 1998.

Oppenheim, M., N. Otani, and C. Ronchi, Hybrid simulations of the saturated Farley-Buneman instability in the ionosphere, Geophys. Res. Lett. 22, 4, 353-356, 1995.

Ortolani, S., Non-linear dynamics and relaxation phenomena in reversed-field-pinch plasmas, in: *Turbulence and Nonlinear Dynamics in MHD Flows, Cargèse, France, July 4-8, 1988*, M. Meneguzzi, A. Pouquet, and P.L. Sulem (editors), Elsevier, Amsterdam, 1989.

Osborne, A.R., and A. Provenzale, Finite correlation dimension for stochastic systems with power-law spectra, Physica D 35, 3, 357-381, 1989.

Packard, N.H., J.P. Crutchfield, J.D. Farmer, R.S. Shaw, Geometry from a time series, Phys. Rev. Lett. 45, 9, 712-716, 1980.

Paulikas, G.A., and J.B. Blake, Effects of the solar wind on magnetospheric dynamics: energetic electrons at the synchronous orbit, in: *Quantitative Modeling of Magnetospheric Processes*, Geophysical Monograph Series, vol. 21, AGU, Washington D.C., 1979.

Pavlos, G.P., G.A. Kyriakou, A.G. Rigas, P.I. Liatsis, P.C. Trochoutsos, A.A. Tsonis, Evidence for strange attractor structures in space plasmas, Ann. Geophys. 10, 5, 309-322, 1992.

Pavlos, G.P., M.A. Athanasiu, D. Kugiumtzis, N. Hatzigeorgiou, A.G. Rigas, E.T. Sarris, Nonlinear analysis of magnetospheric data. Part I. Geometric characteristics of the AE index time series and comparison with nonlinear surrogate data, Non. Proc. Geophys. 6, 1, 51-65, 1999.

Pécseli, H.L., and J. Trulsen, Analytical expressions for conditional averages: a numerical test, Physica Scripta 43, 503-507, 1991.

Pippard, A.B., *Response and Stability: an introduction to the physical theory*, Cambridge University Press, Cambridge, UK, 1986.

Poincaré, H., *Les méthodes nouvelles de la mécanique celeste*, Dover, New, York, NY, 1957.

Preisendorfer, R.W., and C.D. Mobley, *Principal component analysis in meteorology and oceanography*, Elsevier, Amsterdam, Holland, 1988.

Price, C.P., and D. Prichard, The nonlinear response of the magnetosphere – 30 October 1978, Geophys. Res. Lett. 20, 9, 771-774, 1993.

Price, C.P., D. Prichard, and J.E. Bischoff, Nonlinear input-output analysis of the auroral electrojet index, J. Geophys. Res. 99, A7, 13,227-13,238, 1994.

Prichard, D., and C.P. Price, Spurious dimension estimates from time series of geomagnetic indexes, Geophys. Res. Lett. 19, 15, 1623-1626, 1992.

Prichard, D., and C.P. Price, Is the AE index the result of nonlinear dynamics, Geophys. Res. Lett. 20, 24, 2817-2820, 1993.

Press, W.H., B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes in Fortran: The Art of Scientific Computing*, 2^{nd} edition, Cambridge University Press, Cambridge, UK, 1992.

Puente-Baliarda, C., J. Romeu, R. Pous, A. Cardama, On the behavior of the Sierpinski multiband fractal antenna, IEEE Trans. Ant. Prop. 46, 4, 317-524, 1998.

Pulkkinen, T.I., D.N. Baker, L.A. Frank, J.B. Sigwarth, H.J. Opgenoorth, R. Greenwald, E. Friis-Christensen, T. Mukai, R. Nakamura, H. Singer, G.D. Reeves, M. Lester, Two substorm intensifications compared: Onset, expansion, and global consequences, J. Geophys. Res. 103, A1, 15-27, 1998.

Reiff, P., and. J.G. Luhmann, Solar wind control of the polar-cap voltage, in: *Solar Wind-Magnetosphere Coupling*, Y. Kamide and J.A. Slavin (editors), Terra Scientific, Tokyo, 1986.

Ridley, A.J., G. Lu, C.R. Clauer, V.O. Papitashvili, A statistical study of the ionospheric convection response to changing interplanetary magnetic field conditions using the assimilative mapping of ionospheric electrodynamics technique, J. Geophys. Res. 103, A3, 4023-4039, 1998.

Roberts, D.A., D.N. Baker, A.J. Klimas, L.F. Bargatze, Indications of low dimensionality in magnetospheric dynamics, Geophys. Res. Lett. 18, 2, 151-154, 1991.

Ruelle, D., and F. Takens, Nature of turbulence, Comm. Math. Phys. 20, 3, 167-&, 1971.

Ruohoniemi, J.M., and R.A. Greenwald, The response of high-latitude convection to a sudden southward IMF turning, Geophys. Res. Lett. 25, 15, 2913-2916, 1998.

Rugh, W.J., *Linear Systems Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1993.

Russell, C.T., J.G. Luhmann, G. Lu, Nonlinear response of the polar ionosphere to large values of the interplanetary electric field, J. Geophys. Res. 106, A9, 18,495-18,504, 2001.

Samson, J.C., Pure states, polarized waves, and principal components in the spectra of multiple, geophysical time-series, Geophys. J. Roy. Astron. Soc. 72, 3, 647-664, 1983.
Sandahl, I., and E. Jonsson, *AI Applications in Solar-Terrestrial Physics*, Lund, Sweden, 29-31 July 1997, WPP-148, ESA, Noordwijk, Holland, 1998.
Sauer, T., J.A. Yorke, M. Casdagli, Embedology, J. Stat. Phys. 65, 3-4, 579-616, 1991.
Sergeev, V.A., D.G. Mitchell, C.T. Russell, and D.J. Williams, Structure of the tail plasma current sheet at 11 RE and its changes in the course of a substorm, J. Geophys. Res. 98, A10, 17,345-17,365, 1993.
Shan, L.-H., P. Hansen, C.K. Goertz, R.A. Smith, Chaotic appearance of the AE index, Geophys. Res. Lett. 18, 2, 147-150, 1991.
Sharma, A.S., D. Vassiliadis, and K. Papadopoulos, Reconstruction of low-dimensional magnetospheric dynamics by singular spectrum analysis, *Geophys. Res. Lett.* **20**, 335-338, 1993.
Sharma, A.S., M.I. Sitnov, K. Papadopoulos, Substorms as nonequilibrium transitions of the magnetosphere, J. Atmos. Sol.-Terr. Phys. 63, 13, 1399-1406, 2001.
Siscoe, G.L., Solar terrestrial relations – 1984, in: Wolf, R.A., S.-I. Akasofu, S.W.H. Cowley, R.L. McPherron, G. Rostoker, G.L. Siscoe, B.U.O. Sonnerup, Coupling between the solar wind and the Earth's magnetosphere: Summary comments, in: *Solar Wind-Magnetosphere Coupling*, Y. Kamide and J.A. Slavin (editors), Terra Scientific, Tokyo, 1986.
Siscoe, G.L., N.U. Crooker, and K.D. Siebert, Transpolar potential saturation: roles of region 1 current system and solar wind ram pressure, J. Geophys. Res. 107, A10, Art. No. 1321, 2002.
Sitnov, M.I., A.S. Sharma, K. Papadopoulos, D. Vassiliadis, J.A. Valdivia, A.J. Klimas, D.N. Baker, Phase transition-like behavior of the magnetosphere during substorms, J. Geophys. Res. 105, A6, 12,955-12,974, 2000.
Song, P. C.T. Russell, and S.P. Gary, Identification of low-frequency fluctuations in the terrestrial magnetosheath, J. Geophys. Res. 99, A4, 6011-6025, 1994.
Sonnerup, B.U., and L.J. Cahill, Magnetopause structure and attitude from Explorer 12 observations, J. Geophys. Res. 72, 1, 171, 1967.
Stanford Exploration Project, http://sepwww.stanford.edu/public/docs/pvi/toc_html/index.html, 1998.
Stasiewicz, K., Bellan P., Chaston C., Kletzing C., Lysak R., Maggs J., Pokhotelov O., Seyler C., Shukla P., Stenflo L., Streltsov A., Wahlund J.E., Small-scale Alfvenic structure in the aurora, Space Sci. Rev. 92, 3-4, 423-533, 2000.
Stingers, I., and I. Prigogine, *Order out of Chaos: Man's dialogue with Nature*, Bantam, 1989.
Sturrock, P.A., *Plasma Physics: An Introduction to the Theory of Astrophysical, Geophysical, and Laboratory Plasmas*, Cambridge University Press, New York, 1994.
Sun, W., W.Y. Xu, and S.-I. Akasofu, An improved method to deduce the unloading component for magnetospheric substorms, J. Geophys. Res. 105, A6, 13,131-13,140, 2000.
Sutcliffe, P.R., Substorm onset identification using neural networks and Pi2 pulsations, Ann. Geophys. 15, 10, 1257-1264, 1997.
Sutcliffe, P.R., The development of a regional geomagnetic daily variation model using neural networks, Ann. Geophys. 18, 1, 120-128, 2000.
Takalo, J., J. Timonen, and H. Koskinen, Correlation dimension and affinity of AE data and bicolored noise, Geophys. Res. Lett. 20, 15, 1527-1530, 1993.
Takalo, J., J. Timonen, and H. Koskinen, Properties of AE data and bicolored noise, J. Geophys. Res. 99, A7, 13,239-13,249, 1994.
Takalo, J., J. Timonen, Characteristic time scale of auroral electrojet data, Geophys. Res. Lett. 21, 7, 617-620, 1994.
Takalo, J., R. Lohikoski, J. Timonen, Structure function as a tool in AE and Dst time-series analysis, Geophys. Res. Lett. 22, 5, 635-638, 1995.
Takalo, J., and J. Timonen, Neural network prediction of the AE index from the PC index, Phys. Chem. Earth, 24, 1-3, 89-92, 1999.
Takens, F., On the numerical determination of the dimension of an attractor, Lect. Notes Math., 1125, 99-106, 1985.
Taylor, J.B., Relaxation of toroidal plasma and generation of reverse magnetic fields, Phys. Rev. Lett. 33, 19, 1139-1141, 1974.
Ting, A.C., W.H. Matthaeus, D. Montgomery, Turbulent relaxation processes in magnetohydrodynamics, Phys. Fluids 29, 10, 3261-3274, 1986.
Tsonis, A.A., and J.B. Elsner, Beitr. Phys. Atmos. 60, 187-192, 1987.
Tsonis, A.A., A fractal study of dielectric breakdown in the atmosphere, in:Nonlinear Variability in Geophysics, 167-174, D. Schertzer and S. Lovejoy (eds.), Kluwer, 1991.

Tsurutani, B.T., M. Sugiura, T. Iyemori, B.E. Goldstein, W.D. Gonzalez, S.-I. Akasofu, E.J. Smith, The nonlinear response of AE to the IMF B_s driver – a spectral break at 5 hours, Geophys. Res. Lett. 17, 3, 279-282, 1990.
Tsutai, A., C. Mitsui, T. Nagai, Predictions of a geosynchronous electron environment with in situ magnetic field measurements, Earth, Plan. Space 51, 3, 219-233, 1999.
Turcotte, D., Fractals and Chaos in Geology and Geophysics, 2nd edition, Cambridge, 1997.
Uritsky, V.M., and M.I. Pudovkin, Low-frequency 1/f-like fluctuations of the AE index as a possible manifestation of self-organized criticality in the magnetosphere, Ann. Geophys. 16, 12, 1580-1588, 1998.
Uritsky, V.M., A. J. Klimas, D. Vassiliadis, D. Chua, and G. Parks, Scale-free statistics of spatiotemporal auroral emissions as depicted by Polar UVI images: dynamic magnetosphere is an avalanching system, J. Geophys. Res. **107** (A12): art. no. 1426, 10.1029/2001JA000281, 2002.
Valdivia, J.A., G.M. Mlikh, K. Papadopoulos, Red sprites: lightning as a fractal antenna, Geophys. Res. Lett., 24, 24, 3169-3172, 1997.
Valdivia, J.A., G.M. Milikh, K. Papadopoulos, Model of red sprites due to intracloud fractal lightning discharges, Radio Sci. 33, 6, 1655-1668, 1998.
Valdivia, J.A., D. Vassiliadis, A. J. Klimas, A. S. Sharma, and K. Papadopoulos, The spatio-temporal activity of magnetic storms, *J. Geophys. Res.* 104, A6, 12,239-12,250, 1999a.
Valdivia, J.A., D. Vassiliadis, A. J. Klimas, and A. S. Sharma, Modeling the spatial structure of the high-latitude magnetic perturbations and the related current systems, *Phys. Plasmas* 6, 11, 4185-4194, 1999b.
Vassiliadis, D., A. S. Sharma, T. E. Eastman, and K. Papadopoulos, Low-dimensional chaos in magnetospheric activity from AE time series, *Geophys. Res. Lett.* **17**, 1841-1844, 1990.
Vassiliadis, D., A.S. Sharma, K. Papadopoulos, An empirical model relating the auroral geomagnetic activity to the interplanetary magnetic field, Geophys. Res. Lett. 20, 16, 1731-1734, 1993.
Vassiliadis, D., A.J. Klimas, D.N. Baker, D.A. Roberts, A description of the solar wind-magnetosphere coupling based on nonlinear filters, J. Geophys. Res. 100, A3, 3495-3512, 1995.
Vassiliadis, D., and A.J. Klimas, On the uniqueness of linear moving-average filters for the solar wind-auroral geomagnetic activity coupling, J. Geophys. Res., 100, A4, 5637-5641, 1995.
Vassiliadis, D., V. Angelopoulos, D.N. Baker, and A.J. Klimas, The relation between the northern polar cap and auroral electrojet geomagnetic indices in the wintertime, Geophys. Res. Lett. 23, 20, 2781-2784, 1996a.
Vassiliadis, D., A.J. Klimas, D.N. Baker, D.A. Roberts, The nonlinearity of models of the vB(south)-AL coupling, J. Geophys. Res. 101, A9, 19,779-19,787, 1996.
Vassiliadis, D., A.J. Klimas, J.A. Valdivia, D.N. Baker, The D-st geomagnetic response as a function of storm phase and amplitude and the solar wind electric field, J. Geophys. Res. 104, A11, 24,957-24,976, 1999.
Vassiliadis, D., A. J. Klimas, S. G. Kanekal, D. N. Baker, R. S. Weigel, Long-term average, solar-cycle, and seasonal response of magnetospheric energetic electrons to the solar wind speed, *J. Geophys. Res.*, 10.1029/2001JA000506, 2002.
Vecchi, G., D. Labate, and F. Canavero, Fractal approach to lightning radiation on a tortuous channel, Radio Sci. 29, 4, 691-704, 1994.
Vlahos, L., The statistical flare, Ap. J., 1995.
Voros, Z., On multifractality of high-latitude geomagnetic fluctuations, Ann. Geophys. 18, 10, 1273-1282, 2000.
Voros, Z., D. Jankovicova, P. Kovacs, Scaling and singularity characteristics of solar wind and magnetospheric fluctuations, Non. Proc. Geophys. 9, 2, 149-162, 2002.
Weigel, R.S., W. Horton, T. Tajima, T. Detman, Forecasting auroral electrojet activity from solar wind input with neural networks, Geophys. Res. Lett., 26, 10, 1353-1356, 1999.
Weigend, A.S., and N.A. Gershenfeld, *Time Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley, 1994.
Weimer, D.R., S.A. Reinleitner, J.R. Kan, S.-I. Akasofu, Saturation of the auroral electrojet current and the polar-cap potential, J. Geophys. Res. 95, A11, 18,891-18,897, 1990.
Weimer, D.R., Substorm time constants, J. Geophys. Res. 99, A6, 11,005-11,015, 1994.
Wertz, J.R., Space mission analysis and design, 3rd edition, Space Technology Library, Torrance, CA, 1999.
Wu, J.-G., and H. Lundstedt, Prediction of geomagnetic storms from solar wind data using Elman recurrent neural networks, Geophys. Res. Lett. 23, 4, 319-322, 1996.
Wu, J.-G., and H. Lundstedt, Neural network modeling of solar wind magnetosphere interaction, J. Geophys. Res. 102, A7, 14,457-14,466, 1997.
Wu, J.-G., L. Eliasson, H. Lundstedt, A. Hilgers, L. Andersson, O. Norberg, Space environment effects on geostationary spacecraft: Analysis and prediction, Space Weather: Physics and Applications, Adv. Space Res. 26, 1, 31-36, 2000.

Mailing address:
Dimitris Vassiliadis
ST at NASA/Goddard Space Flight Center
Code 612.2
Building 21, Room 265
Greenbelt, MD 20771

E-mail:
vassi@electra.gsfc.nasa.gov

Figures

1. The effect of a particle on the rest of the plasma population extends up to the volume of the Debye sphere. To model dynamics at length scales larger than the sphere one can disregard the degrees of freedom associated with individual particles.
2. Modes as symmetries of the plasma distribution. a. Turbulence is associated with the simultaneous, random activity of many degrees of freedom leading to disorder, a high-level symmetry. b. Symmetry breaking (such as during an instability) produces structures that are embedded in the turbulence.
3. Rugged invariants are preserved at the expense of energy. Energy is lost to (approximately) conserve invariants A_2 and cross-helicity H_c. A significant complication is that there are several paths towards equilibrium (marked as A, B, and C) depending on the initial configuration [Ghosh and Matthaeus, 1990].
4. The lightning discharge is an example of a plasma system with a fractal (self-similar) distribution. Negative of a photograph of a cloud-to-ground stroke showing the discharge branching into progressively finer branches. The shape of the lightning path is determined by the cloud-to-ground potential, water vapor density, and the ground boundary condition.
5. After an image of a dielectric discharge has been digitized, box-counting is applied to determine how much of the space is filled at a spatial resolution Δr [after Kudo, 1998].
6. In an image of a fractal object such as a discharge (cf. Fig. 5), the distribution of the area the object occupies, follows a power law. The power indicates the fractal dimension and degree of self-similarity and space-filling capacity [after Kudo, 1998].
7. Phase space of a particle in one-dimensional potential as a function of the phase space variable q_i and the conjugate momentum p_j.
8. The phase space of an electron close to a Harris sheet has a complex structure. Note the islands, representing regular motion, and the shaded region, indicating unstable, irregular trajectories [Chen and Palmadesso, 1986].
9. The state space of a model of the global dynamics of the terrestrial magnetosphere with emphasis on the substorm cycle. The activity is expressed in terms of the cross-tail electric field and its time derivative. The transition between slow loading and rapid unloading of energy is indicated as the sharp change in the trajectory slope close to E=0.1 [Klimas et al., 1992].
10. The electron-plasma linear oscillator of Eq. (3.1).
11. The power spectrum of the daily solar wind velocity at 1 AU from 1993 to 2000. The peaks at the 27-day solar rotation and its first subharmonic are marked. The first subharmonic of the 11-year solar cycle, at 6.5 years, is also marked.
12. Autocorrelation function of the same daily solar wind velocity data as in Fig. 11. Note the peaks at the 27-day time scale, and its first harmonic and subharmonic.
13. Autocorrelation functions of the three components of the interplanetary magnetic field, as well as its magnitude, |B| (cf. Fig. 12).
14. Principal component ("natural orthogonal component") analysis of the geomagnetic disturbance over the North polar cap and auroral zone: polar view of the ionospheric equivalent-current density in the course of a substorm. Noon is at the top of the figure and dawn at the right [after Sun et al., 2000].
15. Principal component ("natural orthogonal component") analysis of the ionospheric equivalent-current density shown in Fig. 14. The first and second components are shown in the same orientation as in Fig. 14. The first component shows a two-cell pattern, corresponding to the two convection electrojets ("directly-driven disturbance"). The second component shows an intensification of the current in the midnight sector which is interpreted as the contribution of the substorm current wedge ("unloading disturbance") [after Sun et al., 2000].
16. The singular values corresponding to the principal components of high-latitude ionospheric current density [Sun et al., 2000], including the two shown in Fig. 15. Each singular value is the

square root of the covariance in each principal component, which can be interpreted in terms of the magnetic power represented by the component.

17. Calculating the correlation dimension (3.18) of a system's trajectory in a state space is an estimate of the system's degrees of freedom. Ideally the correlation dimension, based on distances from a central point $\mathbf{x}_c$, and more generally pairwise correlations from all points, is independent of the density of points along the trajectory, and represents the dimension of the structure transverse to the trajectory which is of interest. In practice, however, oversampling the system in time as shown here can bias the weight of points along the trajectory (shown here as noisy bands). The bias must be corrected by decimation or other methods [e.g., Theiler, 1986].
18. Effective degrees of freedom in the coupling between solar wind electric field VB_s and geomagnetic index AL as determined from a recurrent neural network model. The performance of the network in producing AL given the interplanetary conditions in VB_s is plotted as a function of the number of hidden nodes in the network [Gleisner and Lundstedt, 2000].
19. a. Schematic of feedback. In a plasma, positive feedback is provided by instabilities. b. Initial condition for two-stream instability. c. Intermediate condition at mode k. d. Saturation.
20. State spaces: a. The phase space of the electron density, $(n_1, dn_1/dt)$. b. A state space constructed from the first two time derivatives of the density is an exact reconstruction of the system's phase space. c. A state space constructed from the first two delays. Here $\tau = 1/2\omega_p$.
21. For an FIR model of two interacting plasmas, the output at time t is the cumulative response to inputs as far back as t-T, where T is the model memory (cf. Eq. (4.23)). Transfer function relating input $\widetilde{B}_z(\omega)$ to output $\widetilde{AL}(\omega)$. a. The power spectrum of the input. b. The power spectrum of the output. c. The ratio between the power spectra is the transfer function of the magnetospheric response [after Tsurutani et al., 1990].
23. Optimization of an ARMA model in terms of model order (state space dimension, m) and nonlinear response (number of nearest neighbors). The model reproduces the AL index from measurements of the $V_{SW}B_s$ solar wind input. The prediction error, Eq. (4.16), is plotted as a function of the two parameters [Vassiliadis et al., 1995a]. The minimum is indicated by an arrow, indicating that the optimal model is nonlinear and of low order for both synthetic and realistic AL index data.
24. A static nonlinear response: transpolar potential saturation [after Siscoe et al., 2002].
25. A dynamic nonlinear response: magnetospheric storage-release dynamics in the course of a substorm, shown here in a numerical simulation by the Lyon-Fedder-Mobarry global-MHD magnetospheric model.
26. The response of the westward electrojet current, represented by the AL geomagnetic index, to the interplanetary input VB_s [Bargatze et al., 1985]. a. Impulse response functions $H(\tau)$ at levels 10 (double-peaks) and 27 (single-peaked) as functions of lag time τ.
27. Stack plot of response functions similar to those shown in Fig. 26, but for 30 activity levels increasing from bottom to top. The directly driven peak at τ=20 min occurs at all activity levels while the loading-unloading peak at τ=60 min appears only at intermediate levels.
28. In a model for the thermodynamics of the magnetospheric plasma sheet [Goertz et al., 1991], the main variables are the cross-tail electric field, the temperature and entropy. The difference between cross-tail field and interplanetary electric field, E_y-E_{sw}, is proportional to the lobe flux, that is, a measure of the energy stored in the system. a. The bifurcation diagram of the plasma sheet dynamics shows the stable levels of the electric field difference E_y-E_{sw} as a function of the activity level. b. The time dependence of the cross-tail field for three levels of interplanetary field E_{SW} driving. Note the qualitative difference of the system activity. The transition from a regular to a chaotic and then to a regular state was modeled to reflect the response of the AL geomagnetic index shown in Fig. 27.
29. Chaos produces an exponential increase in two nearby initial conditions. As a result the prediction error (difference between actual and predicted activity, starting from nearby initial

conditions) increases with a growth rate equal to the sum of positive Lyapunov exponents, or Kolmogorov entropy, K. In order to estimate the sum of positive Lyapunov exponents (the Kolmogorov entropy of Fig. 29) for the AL index, the initial condition is set at an error as small as possible (red curve). In order to estimate the sum of the negative Lyapunov exponents, the initial condition is perturbed and its convergence rate is measured (black curves) [after Vassiliadis et al., 1995a].Nearest-neighbor-based prediction in a state space.