Time series are one of the basic means to extract information for physical systems. [1] Be it in the form of experimental data from a laboratory experiment, observational data for e.g. seismic, astronomical or meteorological events, or data extracted from numerical simulations, time series are usually a starting point in the quest for an adequate description, proper modeling, and eventual prediction of the dynamics of physical systems. Time series can provide information on the underlying dynamics of the system (the deterministic evolution of the velocity of an object in a constant gravitational field, versus the chaotic features of the velocity of car in a sequence of traffic lights) [2], or highlight universalities across seemingly unrelated problems (memory effects in sequences of heartbeats and anomalous diffusion in turbulent magnetic fields) [3-5].

The possibility to extract useful information from time series alone is even more relevant when dealing with physical systems where direct control of experimental variables or numerical parameters is not feasible, which is usually the case in space and astrophysical contexts, resorting to in situ space probe observations and gathering of radiation from remote sources through telescopes.

This is particularly important for the study of space and astrophysical plasmas, which offer a rich variety of phenomena across a wide range of spatial and time scales. Auroral activity and geomagnetic storms as a result of the interaction between the solar wind and the Earth's magnetic field, [6] intermittent behavior in the solar wind, [7] magnetic structures propagating in the interplanetary medium such as coronal mass ejections and magnetic clouds, [8] variations of solar activity as manifested, e.g., in flares and sunspots, [9, 10], stellar variability and formation, [11, 12] or the dynamics of accretion disks around black holes [13], are only a few examples of the variety of physical phenomena where plasma physics is essential. [14-16]

The intrinsically nonlinear features of the plasma interactions pose a major challenge to the understanding of these phenomena, and thus a wide range of approaches have been developed, including observational techniques, analytical models, laboratory experiments, and numerical simulations, which have allowed to understand both basic and highly nontrivial behaviors in plasma systems.

However, nonlinearity in physical systems in general, and plasma physics environments in particular, also leads to the development of complexity through a large variety of manifestations, such as fractal and multifractal dynamics, self-organized criticality, turbulence, or chaos emergence, which are also ubiquitous in space and astrophysical plasmas, [17] and whose study has triggered new techniques and strategies.

The study of complexity in plasmas has provided new ways to describe and understand the nontrivial dynamics in space and astrophysical environments. Certainly, complex behavior is not exclusive to such systems: it is ubiquitous in nature, arising when the microscopic interaction between a large number of individual agents leads to large scale, nontrivial, emergent phenomena. [16, 18-20] It is relevant for the study of problems as diverse as chaos in weather science; many-body gravitational problems; fluid turbulence; collective behavior of ants, fishes, and people; self-organized criticality in sandpiles or earthquakes; fractal patterns in biological systems or electric discharges; virus contagion or opinion propagation, etc.

Therefore, plasma physics has benefited, and can continue to do so, of the various formalisms, approaches, and techniques which have been developed in other, seemingly distant, fields of research.

In this project, we are interested in exploring this idea, by focusing on the study of problems related to space and astrophysical systems, where plasma physics is relevant, and where, due to its nature, time series are the essential source of information available to study the development of complexity features such as fractality, chaos, and turbulence.

Specifically, we propose to do this by means of complex networks, visibility graphs, and fractal approaches. To this end, we will make use of observational data (such as light curves from distant stars, or time-series related to the Sunâs and the Earth's magnetic activities), and data from simulations of magnetohydrodynamic turbulence, which allows to gain intuition on actual physical systems, and carry out systematic studies under controlled conditions.

Complex networks

Complex networks provide a general way to represent systems with a large number of interacting agents. [21] In principle, a complex network is a collection of nodes connected by edges, which can be an abstraction for people and their friendship relationships, cities connected by roads, countries trading goods, or politicians co-sponsoring bills. [22]

The choice of a complex network description for a given system involves, then, a proper definition of the meaning of nodes and of connections between them. These choices depend not only on the system itself, but also on the physics intended to be studied. Thus, various criteria can be used to build a complex network for a given system or data set, which has to be taken into account when interpreting the results obtained.

Various metrics that can be used to characterize the complex network. [23] And, similarly to the choices of nodes and edges, selecting proper metrics is also relevant for an adequate physical interpretation. Metrics can be as simple as the average degree (the average number of connections of nodes in the networks), the average distance (the shortest number of nodes needed to go from one node to another), or the clustering coefficient (related to how nodes group themselves); or more elaborate ones such as centrality (which assess the âimportanceâ of nodes, depending in turn of various criteria such as their relevance to connect regions of the network), vulnerability (measuring the effect of deleting or reconnecting nodes or edges), or community metrics (distinguishing closely connected regions, rather disconnected from the rest of the network), to name a few.

Some of these metrics can measure properties for single nodes (degree, clustering, centrality), which can then be averaged to obtain a representative value for the network as a whole. However, an average may not be a useful value if fluctuations are large, and therefore distributions of these metrics should also be considered to have a better understanding of the network structure.

Given the complexity of the underlying physical systems that we intend to study, it is not expected that a single metric is able to capture its whole richness, and thus a complex network analysis should also consider the use of several metrics, highlighting different features of the relationships between the components that the network represents.

The outcome, however, has proved to be extremely valuable, as complex networks offer new perspectives which are difficult to grasp using more traditional methods. For instance, they can provide information on the type of randomness of the underlying process which leads the system evolution, [24-26] reveal small-world features in networks where the average distance between nodes is much smaller than the number of nodes, [27, 28] and detect hierarchies in the connectivity of nodes leading to nontrivial community structures. [29,30]

Based on the above, complex networks have become an interesting tool for research in a large number of problems. For instance, they have been used to describe the topology of roads connecting cities or points within cities, [31] establish nontrivial correlations in climate dynamics, [32] reveal the structure of rain or wind patterns and their eventual correlation with topographical features, [33, 34] or to study polarization in political systems. [35]

There is also an increasing interest in applying similar ideas to space and astrophysical systems. Complex networks have been used to study the statistical features of solar flare emergence on the solar surface, [36] to establish the correlations of the magnetic field in the auroral region during geomagnetic storms, [37] and to characterize the organization of astrophysical chemical reactions. [38]

These works suggest that complex networks are able to identify nontrivial structures and correlations, and to be a useful approach to follow the dynamics underlying the observed complexity in a variety of systems, including magnetized plasma systems. In particular, complex networks are capable of representing the evolution of spatiotemporal patterns, if nodes are chosen to represent spatial information and edges represent temporal information [39, 40], and this is one of the features that we are interested in further exploring in this project.

Visibility graph

Visibility graphs are another innovative proposal to study time series in a wide variety of fields. [41] The essential idea is that any time series can be mapped into a complex network by using a geometrical criterion of âvisibilityâ between points of the time series. If a physical magnitude is plotted as a function of time, then nodes are defined as the points of the curve, and two nodes are connected if their corresponding points can âseeâ each other. More precisely, two nodes are connected in the network if a straight line can be drawn between them, such that all points between them are below the line. The resulting network is called a visibility graph (VG). We notice that the visibility criterion yields a complex network whose topology is related to the convexity of the curve.

It is straightforward to notice that the resulting network is not modified by uniform vertical or horizontal translations of the curve, nor by uniform stretching of the horizontal or vertical axes, and therefore the visibility graph turns out to be an interesting topological abstraction of the original curve.

To build the visibility graph, lines connecting nodes can be at arbitrary angles, but if the criterion is modified to consider horizontal lines only, then the horizontal visibility graph (HVG) can be defined. In this case, nodes are bars whose height is the value of the variable, and visibility occurs if a horizontal line can be drawn between bars, without being âinterruptedâ by a bar between them. [42]

Normal and horizontal visibility graphs have been shown to discriminate between various kinds of time series, depending on whether they are stochastic, fractal or chaotic, [42] and they are also able to describe the level of irreversibility in the system, by means of metrics such as the Kullback-Leibler distance [43-45]. The visibility graph has also been used for space and astrophysical problems, such as to characterize the sunspots time series which describes solar variability along 11-year cycles, [46] to study solar wind fluctuations during different stages in the solar cycle, [47] and light curves in distant blazars. [48]

We have already used the visibility graph to study time series related to solar and stellar physics (more details below), and we intend to further explore the capability of these graphs to inform us about the statistical and physical properties of these systems.

Fractals and multifractals

Fractals were introduced by Mandelbrot [49] as mathematical objects which have autosimilar properties, that is, zooming in or out them yields the same object. [19] It has turned out to be a useful concept to describe complexity in fields as diverse as geography, biology, or physics.

A typical way to characterize a fractal object is by means of its fractal dimension, which can be a noninteger number. Various definitions are available, such as the box-counting, correlation, and capacity dimensions, [19,50] definitions which, in general, do not yield the same results. If, however, the object can be characterized by a single number, then it is called a monofractal , while multifractals are objects which have a set of fractal dimensions.

Fractals have been shown to be relevant for the description of space and astrophysical systems as well, in many cases due to the multifractality inherent to turbulent systems. [51] For instance, the development of turbulence in the solar wind, and of varying levels of fractality along the solar cycle, [52] or as the distance to the Sun increases, [53] has been a field of active study. The magnetic field structure in the solar photosphere has also been found to exhibit fractal properties, and this has been suggested to be related to the probability of occurrence of solar flares in active regions. [54, 55] Regarding the Earth’s vicinity, turbulence and multifractality have been considered to model the complex behavior of our planet’s magnetosphere. [56]

In previous research, we have studied various aspects of the relationship between fractality and solar and geomagnetic activity. In Ref. [57], the fractal dimension of the Dst index was shown to decrease during geomagnetic storms, a result which was later supported by simulations using the GOY shell model for MHD turbulence [58]. Variations of the intermittency of the shell model with the fractality of its forcing have also been studied, [58] suggesting that the shell model is a useful way to study the response of the magnetosphere under the varying fractal conditions of the interplanetary medium along the solar cycle. [59]

Further work has been done on magnetic time series, by extending the previous work to the use of the SYM-H index, [60] and by showing that the fractal dimension is able to distinguish the stages of magnetic clouds as they move away from the Sun. [61] Fractals have not only been used to study time-series, but also to investigate spatiotemporal patterns such as the evolution of active regions in solar magnetograms along the full 23rd solar cycle. [57]

Based on these works, we can assume that fractality is a useful property to study complex dynamics in space plasmas, which we propose to further study in this project.

Hypothesis

Time-series are one of the basic ways that nature in general, and space and astrophysical systems in particular, conveys information on their underlying dynamics. Complex behavior in plasmas, determined by the nonlinear interactions between fields and charged particles, can be studied by means of approaches developed in other research fields. This not only highlights the universalities inherent to complexity features, but also allows to understand in novel ways the underlying dynamics, complementing other, more traditional approaches, specially in systems of space and astrophysical interest, where time series of in situ or remote observations are the essential way to gather information about them.

Objectives

Based on the discussion above, the hypothesis that complexity in space and astrophysical plasmas can be studied by means of analysis of time series, and also based on our previous works on these subjects, we propose the following general objective:

1. To apply complex networks and fractal/multifractal approaches to various issues related to plasmas physics in space and astrophysical systems.

This can be separated in the following specific objectives:

1. To study, using a complex network approach, the evolution of magnetic structures in the Earth’s magnetosphere and the solar photosphere, in order to determine to what extent the topology of the resulting networks conveys information on solar or geomagnetic activity.

2. To analyze time series related to space and astrophysical systems using the visibility graph approach, in order to take advantage of its capabilities to discriminate the statistical properties of time series, previously reported for simulated data.

3. To use fractal and multifractal methods to study the temporal evolution of complexity in plasma systems, in geomagnetic and solar environments, relating it to the existence of energy dissipation events, and the characterization of structures in magnetic configurations.

Scientific novelty of the proposal

For decades, there has been an increasing realization that complexity is an ubiquitous feature of Nature.

Several approaches have been developed to deal with complexity in physical systems, starting with the traditional ones, observation, experiments, and analytical models, and then simulational approaches as computers have improved their performance during the past century.

However, complexity has posed new challenges, and new techniques such as fractality studies, nonlinear time-series analysis, and complex networks, have shown to be also necessary to properly describe and understand the dynamics leading to complexity in physical systems.

In this project, we intend to advance the understanding of highly nonlinear systems such as plasma environments, by using these techniques, which have already shown to be useful to describe complexity. Complex networks can capture the evolution of spatiotemporal patterns and reveal nontrivial structures in the data, visibility graphs have the potential to discriminate underlying dynamics, and fractals are able to describe the apparently irregular behavior of physical systems. Besides, these methods can extract this information essentially from time series associated to the respective physical processes, which make them particularly useful for space and astrophysical systems, due to the observational restrictions inherent to their study.

Methodology

Given our previous works in subjects related to the ones proposed in this project, we intend to use the methodology described below:

Complex networks We intend to build complex networks in two different ways, in order to capture the dynamics of spatiotemporal patterns.

One strategy is similar to the one we have used previously to study the sequence of sunspots, [40] and which in turn is based on our previous works on seismicity [39,62,63]: each sunspot in a daily magnetogram is regarded as a node, and all nodes in a magnetogram are connected to all nodes in the next one. The resulting networks contain information related to solar activity, information which can be extracted by means of various metrics: degree, clustering coefficient, average distance, betweenness and eigenvector centrality, and their corresponding probability distributions. Since the network is not homogeneous, it has also been interesting to consider the Gini coefficient for each distribution, which yields information on how unequally a certain metric is distributed across the network. [40]

The second strategy is based on correlations between time series. In this case, networks represent the correlations between measurements of a given variable at several places. Nodes are the locations where the instruments are, and two nodes are connected if the similarity between their respective measurements is above a certain threshold. Various similarity criteria can be used (e.g. Pearson correlation, event synchronization [32,64]), and then community detection algorithms [30] can be used to analyze the resulting network. This approach has allowed to identify how perturbations in the magnetic field evolve during geomagnetic storms in the auroral regions, [37] and how rain distribution and wind flow patterns are affected by topographical features. [34, 64]

We intend to use both methods to build complex networks from observed data, in order to understand how spatiotemporal patterns arise and evolve during the nonlinear evolution of space plasmas. In the case of the first strategy, we will improve previous works [59] by adding the tracking of individual active regions by means of image recognition algorithms.

As to the second strategy, involving correlation between time series and community detection algorithms, we propose to apply it to the analysis of the Earth's surface magnetic field fluctuations, and their evolution along varying levels of solar activity along the solar cycle.

Visibility graphs The visibility graph algorithm, both in its normal [41] and horizontal variations [42], will be used to map various time series to complex networks.

Of particular interest to us is its application to time series related to solar activity, as we have already found that comparing two of them (namely, sunspots and mean magnetic field), do not correlate with the solar cycle in the same way, [65] a result which highlights the nontriviality of the mapping, and that results should be carefully analyzed when modifying either the graph (normal vs. horizontal), the time windows analyzed (1-year vs. 11-year windows, for instance), and the metrics considered. The results found in Ref. [65] are being currently extended to other time series, such as solar wind AlfveÌn speed, and solar wind proton flux, in order to establish to which extent these results are general, or particular to the selected time series.

Another interesting use of visibility graphs is the analysis of time series of astrophysical interest. We have already studied light curves of variable pulsating stars taken from the OGLE-III catalog, [66] finding that the VG and HVG have universal features, regardless of the type of pulsation, but results also suggest that some of the metrics are able to discriminate between different pulsation modes. We intend to extend these results by systematically studying the OGLE-III catalog, [67-69] by studying the visible band instead of the radio band as in Ref. [66].

Fractals We will use three methods to associate a fractal dimension to a time series.

One is the box-counting dimension, which has already been used for geomagnetic indices and simulations in Refs. [57, 58, 61, 70]. First, a time window is selected, and a plot of the i-th datum against the i + j-th datum is constructed. j is a certain integer which represents the selected sampling of the time series, and allows to explore the fractality of the series for various timescales. Then, the box-counting fractal dimension [19] for the scatter plot is found.

The above method obtains a dimension by mapping the time series to a bidimensional plot, whose fractal dimension is then calculated in the usual way. Another method, which does not need this mapping into a 2D plot, is the rescaled range analysis (R/S). Here, of the time series divided into windows of varying size, to measure the variability of the series as a function of the window length. This procedure allows to measure the Hurst exponent H of the time series, which in turn can be related to a fractal dimension D if the series is self-similar, by D = 2 â H.

The third method we will consider is also based on 2D plots, and its called the correlation dimension. [71, 72] Essentially, the strategy here involves covering the fractal objects with circles of varying radii (instead of non-overlapping boxes as the box-counting dimension).

The scatter-box counting-box, and the R/S method, will be used to estimate the fractal dimensions of the various time series that we will consider in this project. They are not expected to yield the same results in general, but they both can be good measures of the level of complexity in the time series, by following the evolution of one of these measures as the system evolves, or under different physical conditions.

As for the correlation dimension, it is more suited to 2D plots. We specifically intend to use it to quantify the irregularity of coronal holes boundaries. which has been proposed to be related to the existence of interchange reconnection in the corona. [71] We also plan to compare results with the box-counting method, which can also be used for spatial patterns.

Multifractals If a geometrical object is multifractal rather than monofractal, then it can be better described by a spectrum of fractal dimensions, such as the Renyi fractal dimension spectrum. [72]

We have previously used this strategy to evaluate the multifractality of the set of epicenters and hypocenters for seismic activity in Chile [72], using generalized versions of both the box-counting and correlation dimensions, which yield a family of dimensions D_q , with q a real number. As shown in Ref. [72], though, this method is not reliable for $q < 0$ for experimental data, due to the amplification of experimental errors, and thus we will use regularization techniques based on fits with weighted asymmetric Cantor sets, [73-75] or variations of the method such a the fixed-mass method. [76]

Once the Renyi spectrum is obtained, its Legendre transform yields the singularity spectrum, [73] from which the degree of multifractality can be calculated. [77]

The multifractal analysis will be applied to the various time series and spatial patterns that we have described before. Of particular interest is the estimation of the degree of multifractality, as it has been previously used to characterize the fluctuations of the solar wind as its distance from the Sun increases [78], and also its application to the irregularity of coronal holes boundary, generalizing previous monofractal calculations [71].

Work plan This project is planned to be carried out during 4 years, with the following approximate work plan:

Year 1 Visibility graphs for time series related to solar and stellar activity.

Year 2 Complex network study for sunspots using tracking of active regions. Community structure study of fluctuations of the magnetic field on the Earth’s surface.

Year 3 Fractal study of pulsating stars time series and coronal holes boundaries.

Year 4 Multifractal study of magnetic field time series (Dst and SYM-H geomagnetic indices) and CH boundaries.

Background information to assess the capacity of the team to implement the proposal

In recent years, we have carried out research on various subjects related to the current proposal, and have already used the techniques needed to complete the research plan.

We have explored the usefulness of complex networks for the study of geophysical systems by describing earthquake sequences in the central zone of Chile [39]. This research line has led us to identify useful metrics to describe seismicity, such as the degree distribution and clustering coefficient (which show universal features, regardless of the seismic zone) [39], betweenness centrality (highlighting network topology changes before and after a large event) [63].

In these works, complex networks have been built using the original proposal in Ref. [79], where nodes correspond to locations where seisms occur, and one node A is connected to the node B if the seism at B occurs just after the seism at A.

We have extended this idea to the study of solar daily magnetograms, by considering sunspots as events occurring on the Sunâs surface, and connecting all nodes at a given time to all nodes at the next time. [40]. We have shown that the resulting network carries information about the level of solar activity, but the specific result depends on the chosen metric: some of them correlate with the solar cycle (e.g. degree, decay exponent of the degree distribution), others anticorrelate with it (e.g.eigenvector centrality), whereas other are essentially constant along the cycle (e.g. network density).

Another important issue revealed by this work is the usefulness of the Gini coefficient to analyze the results. Some metrics, such as the clustering coefficient or the betweenness centrality, do not show a clear correlation with the solar cycle, being constant during the solar cycle, but their Gini coefficient do show an interesting behavior. These findings suggest that the method to build the network is useful beyond the original proposal, that the analysis should be based in a variety of metrics in order to have a proper description of the results, and that the Gini coefficient ---a measure usually associated with the distribution of wealth in economic systems--- can also be relevant to space physics.

More recently, we have also studied transport problems in networks, specifically, considering a system of economic agents who interact by trading money. [80] Here, we have considered a system where agents cannot interact with all other agents, and how this may affect the transport and equilibrium (in this case, of wealth) across the network. Similar ideas can be considered in magnetic systems, and we are currently working in collaboration with Dr. Laura Morales and a PhD student, to develop this idea in the context of solar flare occurrence. [81]

Regarding community detection algorithms, we have preliminary results that support the idea that communities within networks of Earth's magnetometers show correlations with solar activity [82], and we intend to investigate this idea further during this project.

We have been involved in research on the use of visibility graphs for the analysis of nonlinear time series. This has been done for sequences of earthquakes in Chile [44], where visibility graphs for the time series of intervals between events, and for seism magnitudes, were considered, and their time reversibility was studied by means of the Kullback-Leibler distance.

We have also applied the VG algorithm to study the light curves of pulsating variable stars, [66] considering various pulsation modes, such as Cepheids, Î´ Scuti stars, stars pulsating in their fundamental or higher harmonic modes, or pulsating in more than one mode simultaneously. We have found some universal behaviors for all stars, regardless of their pulsation modes, and our results also suggest that some metrics can discriminate between stars pulsating in different modes. This is interesting, as it opens the possibility that the VG may be useful to classify some pulsating stars, and we are currently extending this research to a larger star catalog, in order to improve statistics.

We also showed, in Ref. [66], that the VG and the HVG results were essentially unchanged by the presence of observational gaps. Since these gaps are unavoidable given the measurement conditions (e.g., due to changes in the telescope orientation due to the Earth's rotation, or interruptions due to technical issues), these results show that the VG approach may be useful even for incomplete data, which is interesting for the analysis of observational data, instead of experiments or simulations, where conditions and parameters can be better controlled.

Following these ideas, we have also used the VG approach to study two time series related to solar activity, and considering various metrics and time windows: the sunspots number and the Sunâs mean magnetic field [65]. These results confirm the nontriviality of the VG algorithm, and suggest which metrics are more suitable than others, and for what time scales, for the description of solar activity.

As to fractal and multifractal analyses, we have a series of studies on the fractality of geomagnetic time series, [59] such as the Dst time series [57], the SYM-H index (preliminary work in [60]), and the GOY shell model for the simulation of magnetohydrodynamic turbulence [58]. In order to simulate the effect of varying levels of activity in the solar wind influence on the magnetosphere, we have studied the effect of using data obtained from solar wind observations to force the simulation [70, 83].

These results suggest that the box counting fractal dimension is an adequate way to characterize solar activity from the point of view of complexity. This has been extended to the identification of structures in the solar wind such as magnetic clouds [61].

We have also used the correlation dimension to study the distribution of seisms in Chile [72], where the multifractal Renyi spectrum obtained from observed data was extended to negative values of the parameter q, using weighted asymmetric Cantor sets.

Finally, the box counting dimensions for spatial patters has been considered by us to study its correlation with magnetic activity in the Sun along the 23rd solar cycle [57], and image processing algorithms were used to recognize magnetic features on the Sunâs photosphere [40], information which was later used to build complex networks to represent the spatiotemporal pattern on the solar surface.

We believe that the works mentioned, provide us with the necessary experience and intuition to carry out the research proposed for this project.