Publ. Obs. Astron. Belgrade No. 49 (1995), 179-182

ON THE DECIPHERING OF COMPLEX TIME-SERIES

J. NUSPL

Konkoly Observatory, Budapest, P.O.Box 67., H-1525
E–mail nuspl@ogyalla.konkoly.hu

Abstract. The final aim during the deciphering analysis of a recorded data set of time-series observations is to extract the information in the best usable form to understand the dynamical processes working in the background and generating the observable.

In simple cases a frequency decomposition is enough for this purpose but in more complex situations we have to look for more sophisticated methods. A new possibility based on the geometrical picture of phase-space reconstruction methods will also be drafted which can give a more direct bridge between observations and the governing equations.

1. GENERAL FORM OF HARMONICALLY PERTURBED TIME-SERIES

In the first approximation a linear description gives a good fit to the behavior of dynamical systems. According to this the time dependent variations may be considered as harmonic ones and we can use the following mathematical model function

$$f(t) = \sum_i a_i * \sin(\omega_i * t + \phi_i)$$

for the observed time series.

In weakly nonlinear systems where the parameters becomes time dependent we can generally suppose their slow variation. Hence, we can use average values of them for a bit of time interval but the decomposition requires a proper resolution in the frequency domain which means observations long enough in time. However, the two requirements mentioned before can not be satisfied at the same time because increasing the continuous time interval of analysis, the using of quasi constant averages for the parameters will lose its right.

So, in the case of closely spaced frequencies we have to apply another approach if the time dependence is faster.

The time dependence of parameters may be described by a similar decomposition, i.e.

$$a_i(t) = \sum_j A_{ij} * \sin(\alpha_{ij} * t + \sigma_{ij}),$$

$$\omega_i(t) = \sum_j B_{ij} * \sin(\beta_{ij} * t + \chi_{ij}),$$

$$\phi_i(t) = \sum_j C_{ij} * \sin(\gamma_{ij} * t + \psi_{ij}).$$

Of course, the frequencies and phases in these expressions are independent from the basic ones. Inserting these terms into our mathematical model formula the full expression will be

$$f(t) = \sum_i [\sum_j ...] * \sin([\sum_k ...] * t + [\sum_l ...]).$$

This form may be transformed into a simpler one (similar to the original expression) using trivial trigonometric and analytical relations

$$f(t) = \sum_{ijkl} b_{ijkl} * \sin(\Omega_{ijkl} * t + \tau_{ijkl}).$$

The frequencies appeared in the above formula are linear combinations of the original ones

$$\Omega_{ijkl} = \nu_i * \omega_i + n_{ij} * \alpha_j + m_{ik} * \beta_k + r_{il} * \gamma_l$$

and the amplitudes and phases may also be expressed by the original values.

This means, however, that performing the usual Fourier analysis we can get a very complicated and scrambled spectral distribution of the power. Even in the simplest amplitude modulation case, the original form of the function is

$$f(t) = A * \sin(\Omega * t + \Phi) * sin(\omega * t + \phi),$$

and this may be transcribed into

$$f(t) = a_1 * sin((\omega - \Omega) * t + \phi_1) + a_2 * sin((\omega + \Omega) * t + \phi_2).$$

The original frequency disappear and we can detect power in the spectrum at two non-physical frequencies. Of course, the physical interpretation would be completely wrong if one identified these frequencies with real pulsation modes.

One has to note that phase and frequency modulations get similar splitting producing completly mixed power spectra. Nevertheless, we can see from this transformation that the information about the time dependent parameters will be decoded into the pattern of frequencies. Using long term observations we can (in principle) recognize and separate the real frequencies from the detected ones.

2. ARMA PROCESSES AND PREDICTION

We can improve the resolution in frequency using autoregressive AR (or general autoregressive-moving average ARMA) processes. The procedure is iterative and its essence may be understand easily from a geometrical visualization which is equivalent to the generally used one at identification of stochastic processes (see Box, 1971. and Hannan, 1970.).

We consider a dynamical system with several degrees of freedom but we measure only one of the observables which may be even an integrated value, e.g. the total luminosity of a star. The observations distributed equidistantly form a well ordered time series $(m_1, m_2, ..., m_i, ...)$ but we allow gaps in this data set.

This one dimensional observation represents the full underlying dynamical behavior of the system and we can extract that information from it. The followings are based on Takens embedding theorem and its generalizations.

¿From the ordered set of the observations we can form a new set of M-dimensional vectors

$$X^t = (m_t, m_{t-\tau}, ..., m_{t-M*\tau}),$$

where τ means the time interval between two consecutive observations. We can visualize these vectors as points of an M-dimensional space. Due to the theorems mentioned these points will draw up a differentiable manifold topologically equivalent to the original phase-space attractor of the solution. (For example, in the case of a harmonic oscillator measuring the elongation and constructing a 2-dimensional embedding the reconstructed phase-space trajectory will be an ellipse.)

An autoregressive process is defined by

$$X_t = \alpha_1 * X_{t-1} + \alpha_2 * X_{t-2} + ... + \alpha_M * X_{t-M} + \beta_t.$$

The last term of this expression may also be considered as a prediction error if we consider the right hand side as a prediction formula for X_t. In this picture we can describe the AR definition as a linear transformation of the points in the M-dimensional embedding space.

This transformation will drag the points of the reconstructed phase-space trajectory starting from a given observation track. Fitting this dragging path to the reconstructed one we can determine the parameters of the best approximating AR process. During this fitting procedure we have to minimize the cumulative quadratic prediction error for all tracks and data points.

In a second step we can use these parameters to mend the gaps of observation. (Considering the harmonic oscillator the observational tracks will draw up the different segments of the ellipse according to the phase of observation. Using the similar flow of nearby segments we can draw the missing parts of the reconstructed trajectory.) In observational series mended this way can again be analyzed giving a better resolution.

In general we can determine only a local flow of the reconstructed phase-space trajectory and predict forward or backward for a limited interval.

Nonlinear systems may be chaotic whith a diffuse power spectra. In spite of this, the reconstructed phase-space trajectory will tend to fit a simple low dimensional manifold, the attractor of the realized solution (Kolláth, 1990). We can use the local prediction methods described above in this case, too, but we can not decompose it to several independent harmonic components of the variation.

3. FROM TOPOLOGY TO DYNAMICS

In this case the real information about the system is represented by the topological structure of the attractor. However, if our embedding dimension is higher then three, it is difficult or ambiguous (or impossible) task to guess its shape. Hence, we have to use some algorithmic procedure for identification. The algebraic topology supplies these tools in a ready form for us. These methods are widely used in other parts of the physics, e.g. in solid state physics or particle physics.

The topological property of a manifold may be described by its homology group. (The topological features, e.g. twisting, connectedness or unconnectedness, etc., are transformed in this way into an algebraic structure. A good introduction into these technics is given by *Nash*, 1983). The so called 'exact sequences' procedure gives a direct way to determine the homology group of a given manifold.

There exists a one-to-one connection between the homology group of a manifold and the cohomology group of differential forms defined on the manifold. These differential forms correspond to our differential equations used in the every-day work.

Next steps of an algorithmic approach could be given as followings :

- draw the reconstructed phase-space trajectories with different imbedding dimensions
- determine the best dimension from these embeddings
- extract the empirical attractor manifold
- compute and represent the homology group of this manifold
- map it to the dual cohomology group of differential forms

In this way we can get 'directly' the governing equations from our observations.

The application of the procedure mentioned before to real and noisy astrophysical observations is in progress.

References

Box, G. E. P. and Jenkins, G. M. : 1971, "Time series analysis, forecasting and control", Holden Day.

Hannan, E. J. : 1970, "Multiple time series", Wiley.

Kolláth, Z. : 1990, *M.N.R.A.S.* **247**, 377.

Nash, C. and Sen, S, : 1983, "Topology and Geometry for Physicists", Academic Press, New York.