Basic tool for periodic signal detection

Sergio Frasca

28-Dec-1998

Contents

1 Introduction

In this paper the basic techniques of data analysis for the detection of periodic gravitational waves are presented. We start with the easiest model of the signal and the noise and then we introduce some complications that are or can be present in the data. The most promising method for this detection is the power spectrum estimate by the use of the periodograms for the blind search and the lock-in method for the targeted search, with some complication to overcome the problems of the real data. Anyway the real data detection is a very complex problem, up to now largely open (even if in the last years there have been many contributions) and so the detection procedure could be composed of some different stages, at which other procedures can be used; so we also speak briefly of other "non-linear" methods. Because of the introductory character of this paper, also other miscellaneous arguments, that have been of interest in the discussion about the periodic source detection, are presented.

The use of the tecniques here presented is limited by the needed computational power, that grows with high powers of the observing time. So in practice only "hyerarchical" procedures, based on division of the data in pieces, are possible; some of these more advanced procedures will be presented in another paper, *Advanced techniques*.

2 Basic tools

2.1 The basic model

In this section we discuss the basic tools for the detection of small periodic signals in noisy data.

Let us define the more or less simplified hypothesis that we assume:

1. The signal simply adds to the noise, so our data are

$$x(t) = h(t) + n(t) \tag{1}$$

that we know as a set of samples $\{x_i\}$ with sampling time Δt, supposing that such a sampling excludes the aliasing ($1/\Delta t > \nu_{\max}$, being $\nu_{\max}$ the maximum frequency of $x(t)$); $h(t)$ and $n(t)$ are the signal and the noise.

   This hypothesis is not stringent and, in the case of small signals, doesn't require absolute linearity of the detector.

2. The noise is gaussian. Only with this hypothesis some of the results we refer are valid. But in general this is a good approximation of the real gravitational data. The noise is completely described by the power spectrum $S_n(\nu)$; we will use the bilateral form (i.e. defined in the range $-\infty < \nu < \infty$), that, in the case of real (not complex) noise signal, is symmetric about the origin. The total variance of the noise σ_n^2 is the integral in the full range of $S(\nu)$. In the gravitational wave detectors literature the *h-equivalent output noise amplitude density* $H_n(\nu) = \sqrt{S_n(\nu)}$ is normally used.

3. The noise is white, i.e. $S_n(\nu) = const.$ This hipothesis is not too important, because in this simple case the band of interest is very narrow, so it is likely that in the narrow band of interest the noise power spectrum is flat.

4. The noise is stationary. This is an important point, hardly fulfilled in practice. As it is shown in section 3, many types of non-stationarities can be present in the data.

5. The signal is exactly periodic, with angular frequency

$$\omega_0 = 2\pi\nu_0 = \frac{2\pi}{T_0} \tag{2}$$

   and constant amplitude h_0 and phase φ_0. We consider mainly the case of sinusoidal signal

$$h(t) = h_0 \cdot \sin(\omega_0 t + \varphi_0) \tag{3}$$

   but we refer also about methods for non-sinusoidal signals.

6. Absence of "holes" in the data, i.e. the data are continuously observed during the observation period of duration t_{obs}. In practice, this hypothesis can be hardly fulfilled.

7. There are no errors on data timing.

The analysis type will depend on the a priori knowledge. We consider three cases:

1. We know the frequency and the phase of the signal.

2. We know only the frequency.

3. We don't know anything.

In practice the data are sampled, so often we refer to the samples

$$x_i = h_i + n_i \tag{4}$$

We will discuss also the case of non-sinusoidal signal.

2.2 Knowledge of frequency and phase: the *matched filter*

If we know the "exact shape" of the signal and what is unknown is the amplitude h_0 (that is 0 if the signal is not present), as is in the case of knowledge of the frequency and phase for a sinusoidal signal, the optimal detection is performed by means of the **matched filter**. Often the matched filter is applied with known shape and unknown time of occurrence; this, in the case of a sinusoidal signal, means unknown phase: this case, with the name of *cross-correlation filter*, is considered in the next session.

We can compute

$$y(t_{obs)=} \int_0^{t_{obs}} x(t) \cdot \sin(\omega_0 t + \varphi_0) \cdot dt \tag{5}$$

the value of y due to the signal is

$$y_h(t_{obs}) = h_0 \cdot \frac{t_{obs}}{2} \tag{6}$$

the effect of the integration (5) on the noise is to "move" the signal band around zero frequency and apply a low pass filtering that cancels the high frequency components and reduces the variance of the noise to

$$\sigma_n^2 = S_n(\nu_0) \cdot \frac{t_{obs}}{2} \tag{7}$$

so the (quadratic) SNR (signal-to-noise ratio) is

$$SNR_{mf} = \frac{h_0^2 t_{obs}}{2S_n(\nu_0)} = 5 \cdot \left(\frac{h_0}{10^{-26}}\right)^2 \left(\frac{H_n(\nu_0)}{10^{-23} Hz^{-1/2}}\right)^{-2} \left(\frac{t_{obs}}{10^7 s}\right) \tag{8}$$

The distribution of the noise is gaussian. Let us define the parameter k as

$$k = \frac{h_0^2 t_{obs}}{S_n(\nu_0)}, \tag{9}$$

So, in order to have a false alarm probability of 1 %, $k = 4.6$.

2.3 Knowledge of only the frequency: the *lock-in* and the *cross-correlation*

Often the knowledge of the phase of the signal is practically impossible. In this case one can do one of the following procedures:

A) the **two-phases lock-in**[1]. It is performed by computing

$$y(t_{obs}) = \int_0^{t_{obs}} x(t) \cdot \exp(j\omega_0 t) \cdot dt \tag{10}$$

$y(t_{obs})$ is a complex number, of which we consider the square modulus. For the signal,

$$|y_h(t_{obs})|^2 = \left(h_0 \cdot \frac{t_{obs}}{2}\right)^2 \tag{11}$$

so the signal has the same value of the case of the matched filter. The noise is distributed exponentially, with expected value (equal to the standard deviation) $2\sigma_n^2$. The SNR is

$$SNR_{li} = \frac{h_0^2 t_{obs}}{4S_n(\nu_0)} \tag{12}$$

[1]This is not the standard operation of a laboratory lock-in, that has an exponential running integration.

i.e. it is the double of SNR_{mf} ; but there is a loss due to the different distribution of the noise, that is exponential. In order to achieve a false alarm probability of 1 %, we need $k = 18.4$.

B) the **cross-correlation** with the signal $\sin(\omega_0 t)$

$$\widehat{C}\,(\tau) = \frac{1}{t_{obs} - \tau} \int_0^{t_{obs} - \tau} x(t) \cdot \sin[\omega_0 (t + \tau)] \cdot dt \tag{13}$$

$\widehat{C}(\tau)$ is computed for $0 < \tau < \overline{T}$ and it is taken the maximum value. Note that, respect to the case A), the noise variance is one half , but also the mean quadratic value of the signal is one half. This method is practically the same as the matched filter; only the evaluation of the probability is different because of the different *a priori* knowledge.

2.4 Unknown frequency search: *autocorrelation* and *spectrum*

If the frequency of the signal is not known, one can use detection methods based on estimates of the autocorrelation and of the power spectrum (that is the Fourier transform of the autocorrelation).

The method that uses the autocorrelation estimate is heavily based on our simplified hypotheses. The autocorrelation of the data $x(t)$ is equal to the sum

$$R_{xx}(\tau) = R_{hh}(\tau) + R_{nn}(\tau) \tag{14}$$

where

$$R_{hh}(\tau) = \frac{h_0^2}{2} \cos \omega_0 \tau \tag{15}$$

$$R_{nn}(\tau) = \begin{cases} \sigma_n^2 & \text{for} \quad \tau = 0 \\ 0 & \text{for} \quad |\tau| \gg 0 \end{cases} \tag{16}$$

Using an estimation of the autocorrelation for $|\tau| \gg 0$ we have not exactly 0, but a fluctuation value $\varepsilon(t)$ with

$$\sigma_\varepsilon = \sqrt{\frac{1}{B_{eq} t_{obs}} \left[\sigma_n^4 + \sigma_n^2 h_0^2\right]} \simeq \frac{\sigma_n^2}{\sqrt{B_{eq} t_{obs}}} \tag{17}$$

where B_{eq} is the *equivalent bandwidth* of $n(t)$, defined by

$$B_{eq} = \frac{\left(\int_{-\infty}^{\infty} S_n(\nu)d\nu\right)^2}{2\int_{-\infty}^{\infty} S_n^2(\nu)d\nu} = \frac{\sigma_n^4}{2\int_{-\infty}^{\infty} R_n^2(\tau)d\tau} \tag{18}$$

So, in places far from the origin, the SNR is

$$SNR_{ac} = \frac{h_0^2}{2\sigma_n^2}\sqrt{B_{eq}t_{obs}} \tag{19}$$

If the autocorrelation is computed on N samples and the noise samples are uncorrelated, we have

$$SNR = \frac{h_0^2}{2\sigma_n^2}\sqrt{N} \tag{20}$$

Detection can be enhanced of a factor $\sqrt{t_{obs}/T_0}$ taking the average of the $M = t_{obs}/T_0$ periods, because of the uncorrelation of $\varepsilon(t)$; so we have

$$SNR = \frac{h_0^2}{2\sigma_n^2}\sqrt{N \cdot M}$$

The *power spectrum* is the Fourier transform of the autocorrelation. Then an estimation of the power spectrum can be obtained, doing the Fourier transform of the autocorrelation. Another way of doing an estimation of the power spectrum, that gives similar results, is by taking the squared modulus of the Fourier transform of the stretch of data: this is called **periodogram** and it is normally accomplished by a Fast Fourier Transform algorithm, with a strong gain in computation time, that is proportional to $N \log N$ instead of the N^2 of more direct algorithms.

Regarding the periodogram of the $x(t)$ during t_{obs}, the frequency bins have width $\delta\nu = 1/t_{obs}$, so the signal power, that is $h_0 t_{obs}/2$, is divided (in the ideal case) in the two bins at ν_0 and $-\nu_0$. The SNR is the same of 12,

$$SNR_{per} = \frac{h_0^2 t_{obs}}{4S_n} \tag{21}$$

and the distribution is exponential, but, because there are many bins, the probabilistic meaning is different. To discuss a simplified version of the problem, let us consider the case that $S_n(\nu)$ be constant in the band $\Delta\nu$. In that band we have (at least) $\Delta\nu \cdot t_{obs}$ bins (in one day, for a band of 1000 Hz, there are about 10^8 bins). In order to have a false alarm probability of ε, we must have

$$k = 4 \cdot \log \left(\frac{\Delta\nu \cdot t_{obs}}{\epsilon} \right) \tag{22}$$

that, for 10^8 bins and $\varepsilon = 0.01$, is 92.

Another way of performing the estimation in this case, useful in practical cases (as we'll see later), is by dividing the observation time in m intervals, taking the periodograms of the pieces and making the average. In this way we will have

$$SNR_{mper} = \frac{h_0^2 t_{obs}}{4\sqrt{m} S_n} \tag{23}$$

but there are two gains:

- the number of bins is m times lower
- the distribution is no more exponential, but "normalized" χ^2 with $2m$ degrees of freedom.

The value of k to have a false alarm probability ϵ depends on m and on the ratio

$$q = \frac{\epsilon}{\Delta\nu \cdot t_{obs}}$$

The power spectrum estimation by means of periodograms can be enhanced by "windowing" the input chunk of data by particular window functions. This technique reduces the energy that goes in the side lobs of the principal peak. Extensive treatment of the windowing techniques can be found in many introductory texts of signal analysis. The standard approach, anyway, is not sufficient for the real gravitational data analysis, because in that case the frequency of the signal is time-varying and so different optimization procedures must be used.

2.5 Non-sinusoidal signals

Let us suppose that the periodic signal is not simply sinusoidal, as in the case of sources that have the first and the second harmonics of the rotation frequency. If we know the shape, we can perform the matched filter with the shape funcion. Otherwise two techniques can be used:

1. the epoch folding method
2. the multi-harmonics search

With the first method the function

$$y(t) = \sum_{i=0}^{M-1} \left[x(t + i \cdot T_0) - \overline{x} \right] \tag{24}$$

is computed, where $\overline{x}$ is the mean value of $x(t)$, T_0 is the period of the signal and M is the number of periods in t_{obs}. $y(t)$ is defined in the interval $0 \cdot \ t \cdot \ T_0$.

To evaluate the sensitivity of the method, for simplicity, let us suppose that the data are sampled at frequency

$$\nu_s = \frac{N}{T_0} \tag{25}$$

and the noise samples are independent; we have the N values

$$y_i = \frac{1}{M} \sum_{k=1}^{M} \left(x_{i+(k-1)N} - \overline{x} \right) \tag{26}$$

with $M = t_{obs}/T_0$, integer number. Then we can build the variable

$$\chi^2 = \sum_{i=1}^{N} \frac{y_i^2}{\sigma_y^2} \tag{27}$$

and, if the signal $h(t)$ is absent,

$$\sigma_y^2 = \frac{\sigma_n^2}{M} \tag{28}$$

and we can perform a χ^2 test for N degrees of freedom.

If only the signal is present, we should have

$$y_i = h_i - \overline{h} \tag{29}$$

where $\overline{h}$ is negligible (it should be not observable) and we have

$$\chi_h^2 = \frac{M}{\sigma_n^2} \sum_{i=1}^{N} h_i^2$$

This value will be added to the noise, that, in expected value, is N.

If h_i is sinusoidal, as

$$h_i = h_0 \sin \frac{2\pi i}{N} \tag{30}$$

we have

$$\chi_h^2 = \frac{M}{\sigma_n^2} \frac{N h_0^2}{2} \tag{31}$$

With the *multi-harmonics search* method we work in the frequency domain, observing the power spectrum estimate at the m frequencies $\frac{1}{T_0}$, $\frac{2}{T_0}$, $\frac{3}{T_0}$,..., $\frac{m}{T_0}$ and taking the sum of these values (possibly a weighted sum, if there are some a priori information on the harmonics strength or if there is different background noise). In the simplest case this leads to a χ^2 test with $2m$ degrees of freedom.

3 More complex models

In this section we discuss what can be done if one or more of the basic hypothesis were not valid. We refer to cases that are of possible interest in the detection of gravitational waves. The following discussion is essentially qualitative and has just the goal of illustrating the (possible) problems. When the real data will be available, one should recognize the actual problems and then elaborate a detailed strategy.

3.1 Non-linearity

If the interaction between the signal and the noise is not linear, the model can become very complex. If such an interaction were present, one could use it for a better detection, or build a detector in which this phenomenon were enhanced. Anyway this is not foreseen by theory. The smallness of the signal, that can be discriminated from the noise only with very long observations, ensures the correcteness of the linear model.

A different problem arises if the noise has a non-linear dynamics; in this case a better filtering can be achieved by non-linear models (see [1]), but only if it is not gaussian. In fact the optimum filtering theorem, that demonstrates that the optimum filter is linear in the case of additive gaussian noise, excludes any enhancement due to the dynamics of the noise.

3.2 Non-gaussianity

This case can be due to

a) statical non-linearities, as saturations, distorted amplification,...

b) dynamical non-linearities

c) presence of (local) disturbances.

This situation should be corrected or solved at the step of h reconstruction; however a not heavy distortion from gaussianity has no practical impact on the data analysis. If the imput data to a linear filter are not gaussian, generally the output is more gaussian; this because of the central limit theorem and the fact that the linear filter is a linear combination of the input.

3.3 Non-flatness of the noise spectrum

There are two problems associated with this situation:

a) the level of the background noise is slowly varying in the frequency domain: in this case the only problem is that the SNR is different at different frequency; no whitening is needed because the very narrow band of the signal (the noise is supposed to be flat in this narrow band).

b) there are spectral peaks, sometimes slowly changing in frequency: this case, very common in real data, can cause false allarms. Anyway the complex structure of the frequency and amplitude of the signal can be hardly mimicked by a disturbance, but the detection methods must recognize this structure.

3.4 Non-stationarity

The case of non-stationarity is more complex to discuss, because many different types of non-stationarities can be present.

- **Slow variation of the noise statistics**: in this case, in practice, the sensitivity of the detector changes and therefore one must use methods optimezed for this, or, at least, methods that are robust enough for this type of disturbance.In some cases we can decide to analize only data with the minimum value of the background.
- **Pulses or burst disturbances**: these are often wide-band disturbances, that have the effect to increase the power spectrum background noise level in the periodograms. We must subtract them and, if they are not many or very long, it is not important the method (for example one can simply zero the data when a pulse disturbance is present).
- **Varying frequency lines**: if it is possible, we must recognize them.
- **Undetected changes in calibration**: we must determine the amplitude of this effect and, if it is an heavy problem, the use of robust methods (like the Hough transform method) is advisable.

In the stationary case we can define the SNR rate as

$$SNRR = \frac{\partial\, SNR}{\partial\, t_{obs}} = k \frac{h_0^2}{S_n(\nu_0)} \tag{32}$$

where k depends on the particular estimation method. In the non-stationary case we can define

$$SNRR(t) = k \frac{h_0^2(t)}{S_n(t;\nu_0)} \tag{33}$$

and the total SNR is

$$SNR = \int_0^{t_{obs}} SNRR(t)dt \tag{34}$$

3.5 Variable frequency and amplitude signal

The biggest problem in periodic gravitational wave source detection is that the frequency of the signal at the detector is not constant. This happens for various reasons. From the point of view of the signal analysis, there are two main cases:

- The frequency is a stochastic process. This can be the case of emission processes driven by stochastic excitation, as in the case of Wagoner sources, where the signal, in the easiest case, can be modeled as a second order stochastic process.

- The frequency is a deterministic function of time.

The first case is not discussed in this introductory paper. The second one is the case of variations due to

1. Doppler effect due to the motion of the Earth, depending on the position of the source in the sky and the position of the detector on the Earth. This effect has two periodic components, one due to the revolution motion, that has the period of 1 year, a maximum spread of

$$\Delta\nu_{rev} = 0.1986 \cdot \nu_0 \cdot \cos\beta_{ecl} \;\; mHz \tag{35}$$

   where ν_0 is the frequency of the source and β_{ecl} is the ecliptical latitude; the maximum time derivative

$$|\dot{\nu}_{rev}| = \nu_0 \cdot 0.197 \cdot 10^{-10} \cdot \cos\beta_{ecl} \; Hz/s \tag{36}$$

   The other, due to the rotation that has a period of 1 sidereal day, has a maximum spread of

$$\Delta\nu_{rot} = 0.00308 \cdot \nu_0 \cdot \cos\beta_{ter} \cdot \cos\delta \quad mHz \tag{37}$$

   where β_{ter} is the terrestrial latitude of the detector and δ is the declination of thesource; the maximum time derivative of

$$|\dot{\nu}_{rot}| = \nu_0 \cdot 0.112 \cdot 10^{-9} \cdot \cos\beta_{ter} \cdot \cos\delta \;\; Hz/s \tag{38}$$

   The Doppler shift of the observed frequencies can be used as a signature identifying true gw signals and obtain informations on the location of the source.

2. intrinsic causes, as the source not constant rectilinear motion, as in a binary system, or its loss of energy. This can be due to many factors; the most interesting, because it constitutes a lower limit, is the loss of energy caused by the emission of the gravitational waves. In the case of a rotating neutron star, there is a lowering of the frequency that is simply

$$\dot{\nu}_{gw\ damping} = \frac{1}{4}\frac{c^3}{G}\frac{\nu_0 \cdot h_0^2 d^2}{I_{zz}} \tag{39}$$

where d is the distance of the source and I_{zz} is the moment of inertia of the star (typically of the order of $10^{38} \sim 10^{39}\ kg\ m^2$) and h_0 is the amplitude of the wave at the detector.

Although both the effects are deterministic, normally their knowledge is not sufficient to be used for the proposed methods. Small errors in the position of the source in the sky can cause small frequency errors that can cause big phase shifts, and big detection errors, for long enough data sequences. The same happens for small errors in the intrinsic frequency variation. So, the precision needed depends on the length of the data sequence.

Lack of precision (or of knowledge) in the parameters of the source can be overcome by a choice of many "hypothesized" sources, each one with a different position in the sky and/or a different decay parameters for the frequency, in order to cover with sufficient precision the parameters space.

A particular case is that of "all sky search" or blind search, where the hypothesized sources cover densely all the sky. This is a very interesting case, because it is supposed that only a small fraction of the sources are known.

Besides of the variation in frequency, there is also a variation in amplitude, now due to

1. the rotation of the Earth, that changes the angle from which the detector "sees" the source, because of its radiation pattern, that, for a given detector, is a function $G(\alpha, \delta, \psi)$ of the position in the sky (α, δ) and the polarization angle ψ; knowing α, δ, ψ, we can compute the amplitude modulation $g(t)$ for such a source on the given detector. This modulation spreads the power of the signal in side bands mostly at about $\pm 0.116\ mHz$ and $\pm 0.232\ mHz$. Typically the side bands contain about 1/3 of the total received power. Also this effect can be used to identify true gw signal and determine the source position and polarization.

2. intrinsic variation of the structure of the source, that we will neglect here.

These variations are much less problematic than that of the frequency, from the point of view of detection.

If we know the exact frequency variation in time (that means the exact position of the source in the sky and its intrinsic frequency time changes), the signal is

$$h(t) = h_0(t) \cdot \sin\varphi(t) = h_0 \cdot g(t) \cdot \sin\varphi(t) \tag{40}$$

where $g(t)$ is the amplitude modulation and

$$\varphi(t) = \int_0^t \omega(t')dt' + \varphi_0 \tag{41}$$

where $\omega(t)$ is the varying angular frequency of the signal at the detector.

The "classical" method to detect and study a varying frequency signal is by a time-frequency representation, as the Wigner-Ville representation

$$W(t,\omega) = \int x(t-\frac{\tau}{2})x^*(t+\frac{\tau}{2})e^{-j\omega\tau}d\tau \tag{42}$$

The main problem with this method is that the signal-to-noise ratio must be high enough. Better results can be obtained by the use of the matched filter

$$y(t_{obs})= \int_0^{t_{obs}} x(t) \cdot g(t) \cdot \sin(\varphi(t)) \cdot dt \tag{43}$$

if we know the phase of the signal. Otherwise by the lock-in, driven by the varying frequency and modulated by the $g(t)$ in order to "weigh" more the period with higher SNR.

The spectral estimation based on the periodogram, used in the case of unknown non-varying frequency, doesn't work in the case of varying frequency because the energy of the periodic signal is spread on many frequency bins, reducing strongly the signal-to-noise ratio.

If we know, or hypothesize, the position, in the case of unknown source frequency, we can "correct" the Doppler effect (that can be seen as a varying delay in the detection) in the data $x(t)$, obtaining a new signal $x'(t)$ in which the varying frequency sinusoid is transformed in a fixed signal sinusoid, obtained suitably "stretching" the data or suitably "resampling" non-uniformly them. The periodogram of these *resampled data* put all the energy of the periodic signal (if there is no intrinsic variation) in a single frequency bin.

Another serious problem is the possible presence of *glitches*. Glitches are sudden increases of the frequency of the pulsar, that slowly comes back at about the same preceding value. These events are more frequent in the young pulsars, with high $\dot{p}$ (there are no glitches in the millisecond pulsars): in some cases they can occur more times in a year, with duration of the order of ten days each time.

3.6 Presence of holes in the data

Because of the setup operation or presence of disturbances, "holes" could be present in the data. In this case we can take zero as the output of the detector during the holes; this in order to keep the coherence of the data. This is like the data were multiplied by a window function that has normally value 1, and value 0 during the holes.

The presence of holes in the data has two negative effects:

- it reduces the energy of the signal that goes in the detector, reducing the SNR.
- the spectrum estimated by the periodogram is the convolution of the true spectrum and the spectrum of the window, so that the energy of the peaks is partially spread in side bands.

3.7 Detection with more than one antenna

The use of two or more gravitational antennas is of paramount importance for the detection of pulses and of the stochastic background; this is because only the analysis of the coincidences (in the case of the pulses) and of the correlation (in the case of the stochastic background) can exclude local disturbances. In the case of the periodic sources this aspect is less important, because the frequency and amplitude modulation of the signal is very peculiar and, if the SNR is sufficiently high, this excludes the local disturbances.

If we have two or more antennas, we have the following advantages:

- the sensitivity can be enhanced, just summing the output data; in the better case, the quadratic SNR for N antennas is enhanced by a factor N. If the antennas are not in the same place, we must know (or hypothesize) the position of the source in the sky and the sum must be done by suitably delaying the data of the various antennas.

- we can do a cross-spectrum detection.

- we can "confirm" the results of an antenna by the others, diminuishing the chance probability, as it is done with the coincidences of pulse events.

- although normally the detection of a periodic source with an SNR not too low gives the important information on the polarization of the wave, a better work can be done with more antennas, especially in more complex cases (e.g. time variations of the polarization); remember that at a given time, a gravitational antenna "sees" only one polarization, so we need at least two antennas, differently oriented, to have instantaneous informations on the polarization status.

4 Other methods for periodic signal detection

4.1 Introduction

The methods described in section 2 are called "linear methods" because the operations performed on the data are linear operation (with, at most, the square modulus). They have in common the characteristic that the detection SNR (signal-to-noise ratio) to the amplitude of the signal: then they are very good for the case of small signals.

However they have low frequency resolution and create artifacts due to the observation window. Moreover, in order to reduce the estimation error, one must reduce more the resolution.

Non-linear methods have been developed, that overcome, often largely, these limitations; they are also called model-based methods, because they assume a particular model of the data. The limitation is that they work well if the "input" SNR is large enough. For these reasons these methods are not suitable directly for detection of small periodic signals. Nevertheless they are here reported not only for completeness, but also because they could be part of more complex analysis methods.

Here we will present briefly the methods that seems more interesting for the periodic source detection; they are based on the estimation of the autocorrelation of the data $R_{xx}(\tau)$.

4.2 Maximum entropy method (and AR model)

The periodogram power spectrum estimate assumes that the autocorrelation is limited, at most, at the length of the stretch, and it is null outside this range. The *maximum entropy* power spectrum estimate is based on the principle that we know only the first part of the autocorrelation and, outside the known range, the autocorrelation corresponds to that of the most random signal. It was demonstrated that this corresponds to model the data x_i as an autoregressive (AR) process

$$x_i = b_0 u_i + \sum_{k=1}^{m} a_k x_{i-k} \tag{44}$$

where u_i is a white noise sequence with unitary variance and b_0 and the a_k are suitable coefficients. So the method consists in the identification of the coefficients a_k, that can be obtained by the Yule and Walker equations

$$R_{xx}(j) = \sum_{k=1}^{m} a_k R_{xx}(j-k) \tag{45}$$

with $1 \leq j \leq m$, that can be solved in various ways, taking into account that the matrix

$$\mathbf{R} = \begin{pmatrix} R_{xx}(0) & R_{xx}(1) & \dots & R_{xx}(m-1) \\ R_{xx}(1) & R_{xx}(0) & \dots & R_{xx}(m-2) \\ \dots & \dots & \dots & \dots \\ R_{xx}(m-1) & R_{xx}(m-1) & \dots & R_{xx}(0) \end{pmatrix} \tag{46}$$

is a Toeplitz matrix. Then the power spectrum estimation is

$$\widehat{S}_n(\nu) = \frac{b_0^2}{\left|1 - \sum_{k=1}^{m} a_k e^{-j2\pi\nu k}\right|} \tag{47}$$

4.3 Pisarenko method

This method models the data as the sum of m sinusoids and white noise. In this case the autocorrelation is

$$R_{xx}(k) = A_0 \delta(k) + \sum_{j=1}^{m} A_j \cos\left(2\pi k \frac{\nu_j}{\nu_S}\right) \tag{48}$$

$\delta(k)$ is 1 for $k = 0$ and 0 elsewhere. The problem is to find, from the autocorrelation estimate, the values of the A coefficients and the frequencies ν_j.

The AR (autoregressive) model of a sinusoid of frequency ν_j is

$$x_i = b_0 \cdot \delta_{i-i_0} + 2 \cdot \cos\left(2\pi \frac{\nu_j}{\nu_S}\right) \cdot x_{i-1} - x_{i-2} \tag{49}$$

So for m sinusoids the AR model has $2m + 1$ coefficients[2]

$$x_i = b_0 \cdot \delta_i + \sum_{k=1}^{2m} a_k x_{i-k} \tag{50}$$

The solution consists in solving the eigenvector problem

$$\mathbf{R} \cdot \mathbf{a} = \sigma_0^2 \mathbf{a} \tag{51}$$

where $\mathbf{R}$ is the Toeplitz autocovariance matrix estimated from the data, σ_0^2 is the variance of the noise and $\mathbf{a}^T = (1, -a_1, -a_2, ..., -a_{2m})$. From the knowledge of $\mathbf{a}$, we can compute the power spectrum or the amplitude and the frequency of the sinusoids.

5 Miscellaneous problems

5.1 Sampling precision needed

Let us suppose we have data sampled at frequency ν_S, with sampling quantum Δx. Let the data be

$$x_i = h_0 \sin\left(\omega_0 \frac{i}{\nu_S}\right) + n_i \tag{52}$$

The sampling error, due to the value of Δx, is an uncorrelated sequence ε_i (if the signal is small and the noise is white), uniformly distributed in the range $-\Delta x/2 < \varepsilon < \Delta x/2$, with mean value 0 and standard deviation

$$\sigma_\varepsilon = \frac{\Delta x}{\sqrt{12}} \tag{53}$$

[2]For the real data we have 3 degrees of freedom for each sinusoid (the amplitude, the frequency and the phase). In this model the phase is not important and anyway cannot be determined from the autocorrelation. So the degree of freedom are reduced to 2 for each sinusoid and one for the white noise.

The power spectrum of ε_i is then white and

$$S_\epsilon = \frac{\Delta x^2}{12\nu_S} \tag{54}$$

This is a very tiny increase in the background noise n_i

$$\frac{S_\varepsilon}{S_n} = \frac{\Delta x^2}{12\sigma_n^2} \tag{55}$$

For example, with a 16-bit analog-to-digital converter, a good choice can be to take $\Delta x \approx 10^{-3}\sigma_n$ and then the increase in the background noise is only of 1 part in 10^7. So the problem is negligible, if the noise is about white or whitened.

Note that, if the observation time is t_{obs}, we can narrow the bin width of the spectrum to $1/t_{obs}$, so that the signal-to-noise ratio between the sinusoidal signal peak and the sampling noise background is

$$SNR = 6 \cdot \frac{h_0^2}{\Delta x^2} \cdot \nu_S \cdot t_{obs} \tag{56}$$

that, with $\nu_S = 20\,kHz$ and $t_{obs} = 5000\,s$, is 1 for $h_0 \approx 0.4 \cdot 10^{-4}\,\Delta x$. The fact that a sinusoid, sampled with a very big sampling quantum, is well seen in the spectrum, is due to the "benefic" presence of the noise ("dither" effect).

5.2 Subsampling and band extraction

If the data are sampled at a frequency ν_S much higher than that (ν_0) of the signal to search, it can be convenient to reduce in software the sampling frequency of the data to a value ν'_S less than ν_S, but greater than $2 \cdot \nu_0$. This can be done by filtering the data by a *low-pass* filter that makes negligible the power of the data at frequencies over $\nu'_S/2$ (*anti-aliasing filter*) and then sub-sampling the data with frequency ν'_S.

This operation can be performed directly in the frequency domain, with very good results.

This procedure in the frequency domain can be used, with some caution, to extract a band of signal and transfer it at zero Hz; this is the analogous, in the frequency domain, of the heterodyne mechanism that is done in the time domain.

5.3 Log16 format

This is a format that can describe a real number with little more than 16 bits. It best applies to sets of homogeneous numbers. Let us divide the data in sets that are enough homogeneous, as a continuous stretch of sampled data. The conversion procedure computes the minimum and the maximum of the set, checks if the numbers are all positive or negative, or if are all equal, then computes the better way to describe them as a power of a certain base multiplied by a constant (plus a sign). So, any number of the set is represented by

$$x_i = S * m * b^E \tag{57}$$

or, if all the number of the set have the same sign,

$$x_i = m * b^E \tag{58}$$

where

S is the sign (one bit)

m is the minimum absolute value of the numbers in the set

b is the base, computed from the minimum and the maximum absolute value of the numbers of the set

E is the (positive) exponent (fifteen or sixteen bits).

In an header are stored m, b, and a control variables that says if all the number are positive, negative or mixed. For each number two bytes are stored, containing S and E or only E.

The minimum and maximum values can be imposed externally.

The mean percentage error in the case of a gaussian white sequence is better then 10^{-4}.

6 Bibliographic note

A very good introduction to the stochastic processes and probability from the point of view of the signal processing is in [2]. The same author has written a good introduction on continuous and discrete signal theory (see [3]). Other books on the subject are [4], and, more recent, [5] and [6]. A selection of papers of historical (but not only) papers is in [7]

General introductions to signal processing are in [8] [9] and [10], with more applicative issues (but a little dated) in [11]. The first thorough work on the spectral estimation and windowing is in [12].

Good introduction to detection theory are in [13] and [14]; more (and particular) issues are in [15].

An introduction to non-linear time series is in [16]; a good book on the maximum entropy method and other non-linear methods of spectral estimation is [1].

References

[1] Haykin, S., editor. *Nonlinear Methods of Spectral Analysis*. Topics in Applied Physics. Springer-Verlag, (1979).

[2] Papoulis, A. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, third edition edition, (1991).

[3] Papoulis, A. *Signal Analysis*. Mcgraw-Hill, (1977).

[4] Levine, B. *Fondements theoriques de la radiotechnique statistique*, volume I,II,III. Edition de Moscou, (1973).

[5] Hsu, H. *Probability, Random Variables and Random Processes*. Schaum's. McGraw-Hill, (1997).

[6] Viniotis, Y. *Probabilitity and Random Processes for Electrical Engineers*. Mc Graw-Hill, (1998).

[7] Wax, N. *Selected Papers on Noise and Stochastic Processes*. Dover, (1954).

[8] Davenport, W. and Root, W. *An Introduction to the Theory of Random Signal and Noise*. IEEE Press, (1987).

[9] de Coulon, F. *Theory et traitement des signaux*, volume VI of *Traite' d'electricite'*. Presses polytechniques et universitaires romandes, (1998).

[10] Max, J. and Lacoume, J.-L. *Methodes et techniques de traitement du signal et applications aux mesures physiques*, volume vol 1. Masson, 5e edition, (1996).

[11] Max, J. *Methodes et techniques de traitement du signal et applications aux mesures physiques*, volume vol. 2. Masson, 4e edition, (1986).

[12] Blackman, R. and Tukey, J. *The measurement of power spectra*. Dover, (1958).

[13] Wainstein, L. and Zubakov, V. *Extraction of Signal from Noise*. Prentice-Hall, (1962).

[14] McDonough, R. and Whalen, A. *Detection of Signals in Noise*. Academic Press, (1995).

[15] Vaseghi, S. *Advanced Signal Processing and Digital Noise Reduction*. Wiley Teubner, (1996).

[16] Tong, H. *Non-linear Time Series - A Dynamical System Approach*. Oxford Science, (1990).