Astron. Astrophys. 300, 707–710 (1995)

ASTRONOMY AND ASTROPHYSICS

Research Note

On generating power law noise

J. Timmer[1] **and M. König**[2]

[1] Fakultät für Physik, Albert-Ludwigs-Universität, Hermann-Herder-Str. 3, D-79104 Freiburg, Germany
[2] Astronomisches Institut der Universität Tübingen, Waldhäuserstr. 64, D-72076 Tübingen, Germany

Received 12 December 1994 / Accepted 26 January 1995

Abstract. Based on the theory of spectral estimation we propose a new algorithm that is capable to produce the whole variety of possible non deterministic linear time series which exhibit a $(1/f)^\beta$ spectrum. The key point of the new algorithm is to randomize both the phase and the amplitude of the Fourier transform of the data according to its stochastic nature. One possible application is the simulation of AGN lightcurves as well to analyze measured data as to test proposed models.

Key words: X-rays: galaxies – methods: statistical

1. Introduction

A common phenomenon of Active Galactic Nuclei, which presumably harbor supermassive black holes with masses of $10^6 - 10^9 M_\odot$ (Rees 1984), is the strong variability that can be seen in the observed X-ray lightcurves. This variability is often described as flickering or $1/f$ fluctuation (Lawrence et al. 1987). The $1/f$ term describes the distribution of power as a function of frequency in the power spectrum (power density function). A white noise process would generate a curve with constant power in the spectrum, random walk noise would show a $(1/f)^2$ distribution. The presence of $1/f$ fluctuation has also been noted in stellar mass black-hole candidates (Mineshige et al. 1994).

Due to spectral leakage in the Fourier transform – caused by gaps in the observed lightcurve – it is often not possible to derive the characteristics of the present fluctuation from the power spectrum. Therefore, it is necessary to simulate X-ray time series of AGN either to judge the data in a correct manner or to apply other methods to quantify the time variability. A common approach to explain the observed variability is the random superposition of elementary luminosity bursts generated in the accretion processes. If superposed bursts are varied in amplitude and shape a lightcurve similar to $1/f$ fluctuations can be created (Lehto 1989). Such a simulation can only yield an estimate of an actually observed variability. Thus the flickering may not be produced by a superposition of bursts, but by a self organised critical system oscillating at the balance of the accretion rate with radiation pressure (Mineshige et al. 1994). Nevertheless it is very difficult to single out physical features which motivate any particular scenario (Begelman et al. 1991).

One way to generate data that exhibit a power law spectrum $S(f) \sim (1/f)^\beta$ is given in equation (1). In the following the frequency term $\omega = 2\pi f$ will be used.

$$x(t) \sim \sum_\omega \sqrt{S(\omega)} \cos(\omega t - \phi(\omega)) \tag{1}$$

where $\phi(\omega) \in [0, 2\pi]$ is a random phase (Done et al. 1992). Note that this procedure chooses a deterministic amplitude for each frequency and only randomizes the phases. All simulated lightcurves will exhibit a trend caused by the dominating lowest frequency.

In this article we show that the time series produced by this algorithm is only a subset in the set of all possible time series showing the desired spectrum. To do this, we will briefly review some results of the theory of linear stochastic time series in the following section and introduce a new algorithm to generate data with a power law noise in Sect. 3.

2. Mathematical background

There is a fundamental difference between the spectra of periodic, i.e. $x(t+T) = x(t)$, and nonperiodic processes. "Nonperiodic" denotes chaotic as well as (linear and nonlinear) stochastic processes. The spectra of periodic processes show a finite number of peaks, one in the linear case, several higher harmonics in the nonlinear case. The spectra of nonperiodic processes are smooth functions of ω. The behaviour of the spectra of linear stochastic processes is completely understood and some of the results are also applicable to nonlinear stochastic and even to chaotic processes.

In this section we will summarize a main result of the theory of linear stochastic processes. This result is the basis for the new algorithm to generate power law noise which we will present in the next section.

We will start with some definitions:

- The autocovariance function (ACF) is defined as:

$$\mathrm{ACF}(\tau) := < x(t)\, x(t+\tau) > = \lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N-\tau} x(t)\, x(t+\tau) \qquad (2)$$

- The spectrum is defined as the Fourier transform of the ACF:

$$S(\omega) := \sum_{t=1}^{N} ACF(t)\, e^{-i\omega t} \qquad (3)$$

- The periodogram is the squared modulus of the Fourier transform of the data:

$$\mathrm{Per}(\omega) := \frac{1}{N} \left| \sum_{t=1}^{N} x(t)\, e^{-i\omega t} \right|^2 \qquad (4)$$

The notions "spectrum", "periodogram" and "Fourier transform" are often mixed up in the literature. Here, we follow the conventions of Priestley (1989). It should be stressed that the spectrum and the ACF as defined above are quantities that are related to the underlying process and are not related to a realization of the process. On the other hand, the periodogram is related to each new realization and the relation between the spectrum and the periodogram is the important point to be discussed now.

We will start this discussion with the white noise process and outline the straightforward generalizations to non trivial processes afterwards.

A normal distributed white noise process

$$x(t) = \epsilon(t), \quad \epsilon(t) \sim \mathcal{N}(0, \sigma^2) \qquad (5)$$

with mean zero and variance σ^2 is characterized by the ACF:

$$\mathrm{ACF}(\tau) = \sigma^2\, \delta(\tau) \qquad (6)$$

Thus, the spectrum as the Fourier transform of the δ function is a constant:

$$S(\omega) = \sigma^2 \qquad (7)$$

The Fourier transform $f(\omega)$ of a realization of this process written as real and imaginary part is given by

$$f(\omega) = \frac{1}{\sqrt{N}} \sum_t x(t)\, \cos(\omega t) + i\, \frac{1}{\sqrt{N}} \sum_t x(t)\, \sin(\omega t) \qquad (8)$$

A main result of the theory of spectral estimation (Priestley 1989) is that $f(\omega)$ is a complex gaussian random variable

$$f(\omega) = \mathcal{N}(0, \frac{1}{2} S(\omega)) + i\, \mathcal{N}(0, \frac{1}{2} S(\omega)) \quad , \qquad (9)$$

whose variance does not depend on the number of data points. These random variables are uncorrelated for different frequencies:

$$< f(\omega_i)\, f(\omega_j) > = \mathrm{const}\ \delta_{ij} \qquad (10)$$

The periodogram is given by:

$$\begin{aligned} \mathrm{Per}(\omega) &= |f(\omega)|^2 \\ &= \frac{1}{N} \left(\sum_t x(t)\, \cos(\omega t) \right)^2 + \\ &\quad \frac{1}{N} \left(\sum_t x(t)\, \sin(\omega t) \right)^2 \end{aligned} \qquad (11)$$

and as the sum of two squared gaussian distributions follows a χ^2 distribution with two degrees of freedom χ^2_2:

$$\mathrm{Per}(\omega) \sim \frac{1}{2} S(\omega) \chi^2_2 \qquad (12)$$

again independent of N.

Since the mean and the variance of χ^2_2 are two and four respectively, the standard deviation of the periodogram is equal to the mean, i.e.

$$\mathrm{Per}(\omega) = S(\omega) \pm S(\omega) \quad . \qquad (13)$$

Thus the periodogram is fluctuating wildly and its variance is independent of N, the number of data points, $\mathrm{Per}(\omega)$ is not a consistent estimator of the spectrum since its variance does not decrease with N. These results not only hold for linear stochastic processes but also for nonlinear stochastic and even for most chaotic processes.

For linear stochastic processes the spectrum and the periodogram are obtained by multiplying the results for the white noise process by the filter function – here: a power law – that describes the process. For these processes the variance of the complex random variable $f(\omega)$ in Eq. (9) becomes frequency dependent and is determined by the spectrum. Equation (13) remains valid. This is also true for nonlinear stochastic processes and even for most chaotic processes. For these processes Eq. (10), i.e. the orthogonality of different Fourier components, does not hold in general.

To summarize the results with respect to the simulation problem of power law noise:

The standard method of generating these time series according to Eq. (1) reflects only one part of the stochasticity of the Fourier transform of nonperiododic processes, namely the randomness of the phases. Choosing the amplitudes equal to the square root of the spectrum, it does not take into account the randomness of the periodogram according to the χ^2_2 distribution. In order to create power law time series it is necessary to allow randomness both in phases and in amplitudes.

3. A new algorithm to simulate power law noise

The new algorithm is based on Eq. (9), which connects the desired spectrum with the variance of the complex random variable $f(\omega)$.

The algorithm is defined by the following steps:

- Choose a power law spectrum $S(\omega) \sim (1/\omega)^{\beta}$.

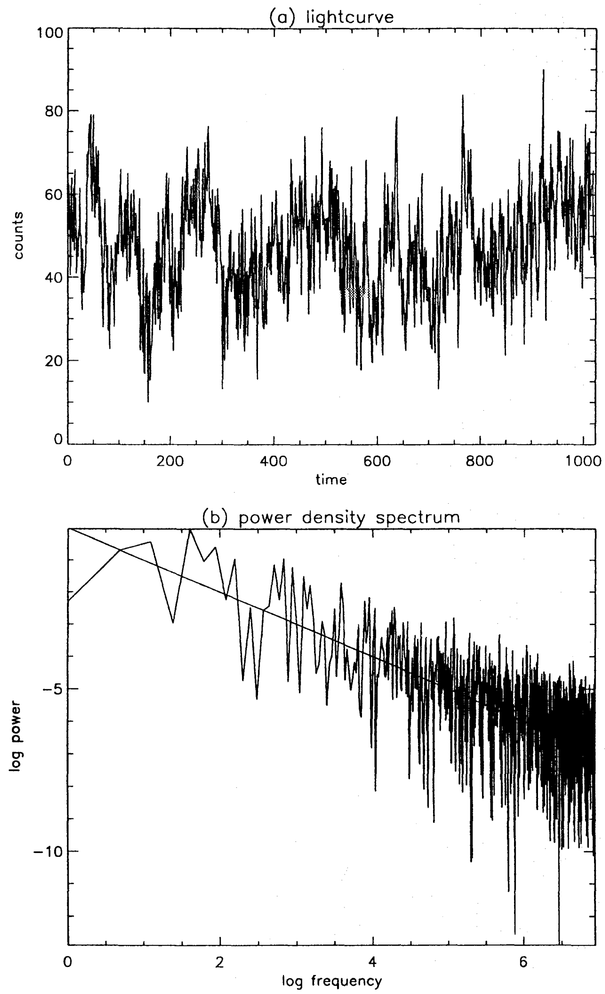

Fig. 1. a Simulated flicker noise lightcurve (β=1.0, N=1024). **b** Corresponding spectrum and periodogram. The time and counts units are arbitrary

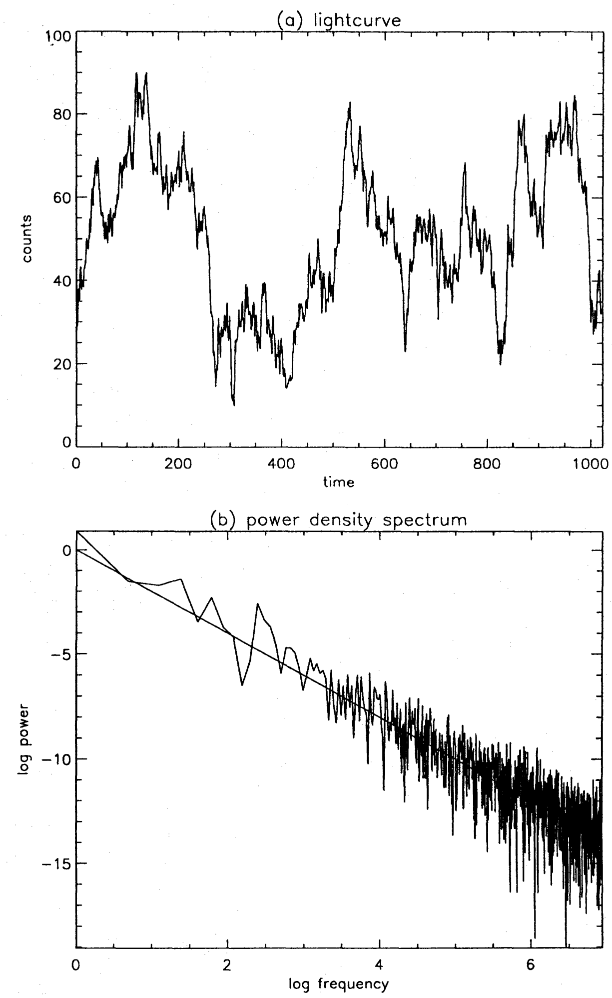

Fig. 2. a Simulated random walk noise lightcurve (β=2.0, N=1024). **b** Corresponding spectrum and periodogram. The time and counts units are arbitrary

- For each Fourier frequency ω_i, draw two gaussian distributed random numbers, multiply them by $\sqrt{\frac{1}{2}S(\omega_i)} \sim (1/\omega)^{\beta/2}$ and use the result as the real and imaginary part of the Fourier transform of the desired data.
- In the case of an even number of data points, for reason of symmetry $f(\omega_{Nyquist})$ is always real. Thus only one gaussian distributed random number has to be drawn.
- To obtain a real valued time series, choose the Fourier components for the negative frequencies according to $f(-\omega_i) = f^*(\omega_i)$ where the asterisk denotes complex conjugation.
- Obtain the time series by backward Fourier transformation of $f(\omega)$ from the frequency domain to the time domain.

Due to the fact that Fast Fourier Techniques can be used to evaluate the lightcurve, this new algorithm is even faster than the deterministic method described in Eq. (1).

4. Discussion

Using this algorithm the full variety of possible time series showing the same spectrum can be explored. Especially, in case of flicker noise (Fig. 1) where the first frequency bin contributes the largest part of the variance, this algorithm ensures that the first frequency bin does not dominate the time series in a deterministic manner, but according to its natural fluctuations.

Choosing β=2.0 for the simulation, a random walk lightcurve is generated (Fig. 2). Compared to the flicker noise lightcurve such a random walk curve is dominated by longer timescales. Figure 3 presents a real X-ray lightcurve of the Seyfert Galaxy NGC 5506 as observed by EXOSAT in 1986. The corresponding spectrum shows a slope of 1.8±0.3 (Lawrence & Papadakis 1993) indicating the numerical relation to random walk noise. In contrast to this the naked eye would classify the observed variability more likely as flicker noise (McHardy & Czerny 1987). This reveals the main problem of estimating the slope of the spectrum: Its evaluation depends on the chosen frequency regime in which the fit of the slope fit is

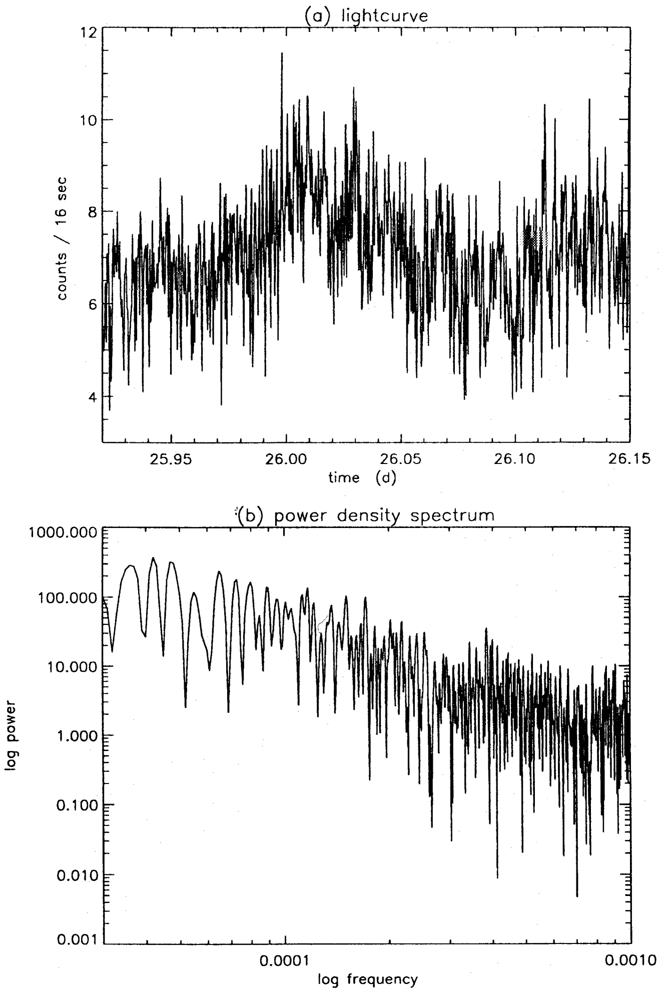

Fig. 3. **a** X-ray lightcure (detail) of NGC 5506 (EXOSAT-ME Jan./1986). **b** Corresponding periodogram. The flattening at low frequencies is here due to the finite observation length

done. According to this, the algorithm is a good tool to judge the method used to compute the slope of the spectrum.

Another application of the presented algorithm is the proper estimation of 1σ–errors by Monte Carlo simulations in period searches using epoch folding techniques. Using Eq. (1) to generate the non deterministic part of a periodic lightcurve the 1σ–error of the related distribution of the estimated periods is underestimated significantly. In the case of the cataclysmic variable RXJ1940.1-1025 the period search employing simulated data with Eq. (1) yields an error of ± 3 sec (Done et al. 1992). If the new algorithm is applied, the resulting 1σ–error is about 5 to 10 times larger.

References

Begelman, M.C., De Krool, M. 1991, in Variability in Active Galactic Nuclei, ed. H.R.Miller & P.J.Wiita (Cambridge: Cambridge Univ.Press), 198
Done, C., Madejski, G.M., Mushotzky, R.F. et al. 1989, ApJ 400, 138
Lawrence, A., Watson, M.G., Pounds, K.A. et al. 1987, Nat 325, 694
Lawrence, A., Papadakis, P. 1993, ApJS 414, 85
Lehto, H.J. 1989, in Proc. 23d ESLAB Symp. on Two Topics in X-ray Astronomy, Vol.1, ed. J.Hunt & B.Battrick (ESA SP-296), 499
McHardy, I., Czerny, B. 1987, Nat 325, 696
Mineshige, S., Ouchi, N.B., Nishimori, H. et al. 1994, PASJ 46, 97
Priestley, M.B. 1989, Spectral Estimation and Time Series, (San Diego: Academic Press), 389
Rees, M.J. 1984, ARA&A 22, 471

This article was processed by the author using Springer-Verlag $\LaTeX$ A&A style file version 3.