

# Time-Frequency Analysis

## — $G(\lambda)$ -stationary Processes

Huiping Jiang, Henry L. Gray and Wayne A. Woodward

April 30, 2004

### **Abstract**

Methods such as wavelets and M-stationary process have been developed to analyze the time-frequency properties of a process where frequency changes with time. In certain circumstances, when the frequencies of a process change systematically either monotonically increasing or monotonically decreasing across time, another approach is to apply an appropriate Box-Cox transformation to the time axis for the given signal in order to obtain a new stationary data set. This new data set can be analyzed by standard methods. Processes which are transformed to a stationary process after Box-Cox transformation on the time scale are called  $G(\lambda)$ -stationary processes, where  $\lambda$  is the corresponding parameter of the Box-Cox transformation. The method is illustrated with analysis of both simulated and real data. Finally, it is shown that such processes can be transformed to stationarity by sampling properly.

# 1 Introduction

Time varying frequencies are quite common in speech, biological data, geophysical processes and so on. Traditional Fourier analysis under the assumption of stationarity may not be applicable to these processes since the frequency content is evolving over time. The idea of time deformation was introduced to transform the time axis in order to change these non-stationary processes to stationary processes. Stock (1987, 1988) formalized a time deformation model and captured cyclical behavior of some macroeconomic data by using the deformed time modeling, that was not detected in the original equally spaced calendar time modeling. Meanwhile, Gray and Zhang (1988) developed continuous multiplicative-stationary(M-stationary) processes for the purpose of analyzing non-stationary data with cyclical behavior changing approximately linearly in time. A continuous M-stationary process can be transformed to a continuous weakly stationary process, which is referred to as the dual process, through a logarithmic time transformation. One type of M-stationary process, the continuous Euler process, whose corresponding dual is the continuous autoregressive(AR) processes, was proposed by Gray and Zhang (1988). Since these processes are continuous, the model was never applied to data even though the process has remained of some theoretical interest. Gray, Vijverberg and Woodward (2004) extended the M-stationary processes to the case of discrete data and developed the corresponding discrete Euler process. Choi (2003) defined the continuous and discrete mixed M-stationary Euler( $p, q$ ) processes, whose duals are the usual continuous and discrete ARMA processes respectively. Compared to the Euler model, the Euler( $p, q$ ) model typically has fewer parameters.

In Gray, Vijverberg and Woodward (2004), it was shown that when the period of a process is increasing approximately linearly with time and the process has clear

cyclic behavior, M-stationary models give a better fit from both the spectral analysis and forecast performance points of view than traditional methods.

The objective of the current article is to apply the Box-Cox transformation to the time axis and then introduce a new class of processes,  $G(\lambda)$ -stationary processes, based on this transformation. This class is quite flexible and can represent processes with a wide range of time-varying frequency behavior. The stationary processes and the M-stationary process are simply two special cases of the  $G(\lambda)$ -stationary process, i.e., the cases  $\lambda = 1$  and  $\lambda = 0$ , respectively.

The outline of this article is as follows. The  $G(\lambda)$ -stationary process will be introduced in Section 2. The instantaneous period and instantaneous frequency of the  $G(\lambda)$ -stationary process will be discussed in Section 3. The  $G(\lambda)$ -autoregressive-moving average ( $G(p, q; \lambda)$ ) process will be introduced in Section 4, and a sampling method for continuous  $G(\lambda)$ -stationary processes will be discussed in Section 5. In Section 6, the instantaneous spectrum of  $G(\lambda)$ -stationary processes is defined. The origin problem and equally spaced sampling from  $G(p, q; \lambda)$  processes will be discussed in Section 7. In Section 8, we discuss the estimation of the parameter,  $\lambda$ , of the Box-Cox transformation and the origin offset of the data,  $\Lambda$ . Two application examples will be shown in Section 9; and in Section 10 we give concluding remarks.

## 2 $G(\lambda)$ -Stationary Processes

Hannan (1965) considered the concept of stationarity under a general group composition law, i.e.  $E((X(t) - \mu)(X(t \circ \tau) - \mu)) = C_X(\tau)$ , where “ $\circ$ ” denotes a group composition law. If we let  $f(t, \tau) = t \circ \tau$ , this general stationarity can be written as  $E((X(t) - \mu)(X(f(t, \tau)) - \mu)) = C_X(\tau)$ . Different group composition laws re-

late to different types of time-changing frequency or periodic behavior. The usual notion of stationarity, of course, assumes that the group composition law is addition, which corresponds to the processes with fixed periods and frequencies over time. For the M-stationary process, the composition law is multiplication, which results in periodic behavior that changes linearly in time. Given the general stationary process, the time deformation technique proposed here attempts to find the transformation function,  $g$ , between the general group composition law and the usual additive composition law, i.e.,  $g(f(t, \tau)) = g_1(t) + g_2(\tau)$ , where  $g_1$  and  $g_2$  are two functions. For the M-stationary process,  $f(t, \tau) = t\tau$  and  $g(t) = g_1(t) = g_2(t) = \ln(t)$ . The  $G(\lambda)$ -stationarity proposed here is based on another group composition law,  $f(t, \tau) = (t^\lambda + \tau\lambda)^{\frac{1}{\lambda}}$ , where  $\lambda \in (-\infty, \infty)$ . We use the Box-Cox transformation function  $g$ , defined by  $g(t) = g_1(t) = \frac{t^\lambda - 1}{\lambda}$ , and  $g_2(\tau) = \tau$ . The definition follows.

**Definition 2.1** *Let  $X(t)$  be a stochastic process defined for  $t \in (0, \infty)$  such that for any  $(t^\lambda + \tau\lambda) \in (0, \infty)$ , and constant  $\lambda \in (-\infty, \infty)$ ,*

$$(i) \quad E[X(t)] = \mu,$$

$$(ii) \quad \text{var}[X(t)] = \sigma^2 < \infty,$$

$$(iii) \quad E[(X(t) - \mu)(X((t^\lambda + \tau\lambda)^{\frac{1}{\lambda}}) - \mu)] = B_X(\tau; \lambda).$$

*Then  $X(t)$  will be called a  $G(\lambda)$ -stationary process.*

Remark: throughout this paper, “stationary” means “weakly stationary”, but we take the word “weakly” as understood.  $B_X(\tau; \lambda)$  here will be referred to as the  $G(\lambda)$ -autocovariance so that  $\text{var}(X(t)) = B_X(0; \lambda)$ . The  $G(\lambda)$ -autocorrelation is defined by

$$\rho_X(\tau; \lambda) = \frac{B_X(\tau; \lambda)}{\text{var}(X(t))}.$$

When  $\lambda = 1$ ,  $B_X(\tau; \lambda) = E[(X(t) - \mu)(X(t + \tau) - \mu)]$ . Note that  $X(t)$  here has the covariance properties associated with a stationary process in the usual sense. The only difference is that  $X(t)$  is restricted to  $t > 0$ . In practice, we only observe the stationary process at  $t > 0$ , i.e.,  $t \in (0, \infty)$ . Then a  $G(\lambda)$ -stationary process with  $\lambda = 1$  can be considered to be an “observable stationary process.” Since

$$\lim_{\lambda \rightarrow 0} B_X(\tau; \lambda) = E[(X(t) - \mu)(X(te^\tau) - \mu)] = R_X(e^\tau) = R_X(\varsigma),$$

where  $\varsigma = e^\tau$ , the limiting case  $\lambda = 0$  corresponds to an M-stationary process. We will refer to the limiting case as  $\lambda = 0$ . The usual autocovariance is

$$C_X(h; t) = E[(X(t) - \mu)(X(t + h) - \mu)] = B_X\left(\frac{(t + h)^\lambda - t^\lambda}{\lambda}; \lambda\right).$$

So the autocovariance of a  $G(\lambda)$ -stationary process depends on both time and lag when  $\lambda \neq 1$ . Therefore, the  $G(\lambda)$ -stationary process with  $\lambda \neq 1$  is non-stationary in the usual sense.

**Definition 2.2** *Let  $Y(u)$  be a stationary stochastic process over  $(-\infty, \infty)$ . Let  $u = g(t)$  and  $X(t) = Y(u)$  on  $t \in (0, \infty)$ . Then  $\{Y(u); u \in (-\infty, \infty)\}$  will be called the stationary dual process of  $\{X(t); t \in (0, \infty)\}$ .*

For a  $G(\lambda)$ -stationary process, we have  $u = g(t) = \frac{t^\lambda - 1}{\lambda}$ , and the dual process is a stationary process  $Y(u)$  such that  $Y(u) = X(t)$  on  $t \in (0, \infty)$ , i.e., for

$$u \in \begin{cases} (-\frac{1}{\lambda}, \infty) & \text{if } \lambda \geq 0 \\ (-\infty, -\frac{1}{\lambda}) & \text{if } \lambda < 0 \end{cases}.$$

The function  $g(t)$  will be referred to as the transformation function.

**Theorem 2.1**  *$X(t)$ ,  $t \in (0, \infty)$ , is  $G(\lambda)$ -stationary if and only if it has a stationary dual  $Y(u)$  with the transformation function  $u = g(t) = \frac{t^\lambda - 1}{\lambda}$ .*

Proof: Clearly, properties (i) and (ii) concerning constant mean and constant finite variance hold for  $X(t)$  if and only if they hold for  $Y(u)$ .

Since  $u = \frac{t^\lambda - 1}{\lambda}$ , then  $t = (u\lambda + 1)^{\frac{1}{\lambda}}$  and therefore

$$t \in (0, \infty) \iff \begin{cases} u \in (-\frac{1}{\lambda}, \infty) & \text{if } \lambda \geq 0, \\ u \in (-\infty, -\frac{1}{\lambda}) & \text{if } \lambda < 0. \end{cases}$$

$\implies$  If  $X(t)$  is  $G(\lambda)$ -stationary and  $B_X(\tau; \lambda)$  is the  $G(\lambda)$ -autocovariance, then  $B_X(\tau; \lambda) = E[(X(t) - \mu)(X((t^\lambda + \tau\lambda)^{\frac{1}{\lambda}}) - \mu)]$ , where  $t > 0$  and  $t^\lambda + \tau\lambda > 0$ . Let  $Y(u) = X(t)$ .

If  $\lambda = 0$ , we have  $u = \ln(t)$  and

$$\begin{aligned} C_Y(\tau) &= E[(Y(u) - \mu)(Y(u + \tau) - \mu)] \\ &= E[(X(e^u) - \mu)(X(e^{u+\tau}) - \mu)] \\ &= E[(X(e^u) - \mu)(X(e^u e^\tau) - \mu)] \\ &= R_X(e^\tau), \end{aligned}$$

where  $u \in (-\infty, \infty)$  and  $\tau \in (-\infty, \infty)$ . Since  $X(t)$  is M-stationary,  $C_Y(\tau) = R_X(e^\tau)$  does not depend on  $t$ , and  $Y(u)$  is the stationary dual of  $X(t)$ . If  $\lambda > 0$ , letting  $Y(u) = X(t)$ , where  $u = \frac{t^\lambda - 1}{\lambda}$ , we have

$$\begin{aligned} C_Y(\tau) &= E[(Y(u) - \mu)(Y(u + \tau) - \mu)] \\ &= E[(X(t) - \mu)(X((t^\lambda + \tau\lambda)^{\frac{1}{\lambda}}) - \mu)] \\ &= B_X(\tau; \lambda), \end{aligned}$$

where  $u \in (-1/\lambda, \infty)$  and  $\tau \in (-1/\lambda - u, \infty)$ . Then  $Y(u); u \in (-\infty, \infty)$  is the stationary dual of  $X(t)$ . The results for  $\lambda < 0$  are similar.

$\Leftarrow$  If  $Y(u)$  is stationary and  $C_Y(\tau)$  is the corresponding autocovariance, then for

any  $t > 0$  and  $t^\lambda + \tau\lambda > 0$ , we have

$$\begin{aligned}
B_X(\tau; \lambda) &= E[(X(t) - \mu)(X((t^\lambda + \tau\lambda)^{\frac{1}{\lambda}}) - \mu)] \\
&= E[(Y(\frac{t^\lambda - 1}{\lambda}) - \mu)(Y(\frac{t^\lambda - 1 + \tau\lambda}{\lambda}) - \mu)] \\
&= E[(Y(\frac{t^\lambda - 1}{\lambda}) - \mu)(Y(\frac{t^\lambda - 1}{\lambda} + \tau) - \mu)] \\
&= C_Y(\tau).
\end{aligned}$$

Hence,  $B_X(\tau; \lambda)$  only depends on  $\tau$ . Therefore  $X(t)$  is a  $G(\lambda)$ -stationary process.

**Property 2.1** *Let  $X(t)$  be a  $G(\lambda)$ -stationary process, and let  $Y(u)$  be its dual. The following hold:*

$$(i) \quad B_X(\tau; \lambda) = \begin{cases} C_X(\tau) & \lambda = 1 \\ R_X(e^\tau) & \lambda \rightarrow 0 \end{cases}$$

(ii)  $B_X(-\tau; \lambda) = B_X(\tau; \lambda)$  for all values of  $\lambda$ .

Proof:

(i) If  $\lambda = 1$ ,

$$\begin{aligned}
B_X(\tau; \lambda) &= E[(X(t) - \mu)(X((t^\lambda + \tau\lambda)^{\frac{1}{\lambda}}) - \mu)] \\
&= E[(X(t) - \mu)(X(t + \tau) - \mu)] \\
&= C_X(\tau),
\end{aligned}$$

and thus in this case, the  $G(\lambda)$ -autocovariance is the usual autocovariance function.

If  $\lambda \rightarrow 0$ , under condition that  $C_X(h; t)$  is continuous for any  $t > 0$ ,

$$\begin{aligned}
\lim_{\lambda \rightarrow 0} B_X(\tau; \lambda) &= \lim_{\lambda \rightarrow 0} E[(X(t) - \mu)(X((t^\lambda + \tau\lambda)^{\frac{1}{\lambda}}) - \mu)] \\
&= E[(X(t) - \mu)(X(\lim_{\lambda \rightarrow 0} (t^\lambda + \tau\lambda)^{\frac{1}{\lambda}}) - \mu)] \\
&= E[(X(t) - \mu)(X(te^\tau) - \mu)] \\
&= R_X(e^\tau),
\end{aligned}$$

and so the  $G(\lambda)$ -autocovariance is the M-autocovariance.

(ii) Letting  $\tau > 0$  and  $s = (t^\lambda - \tau\lambda)^{1/\lambda} > 0$ , then

$$\begin{aligned} B_X(-\tau; \lambda) &= E[(X(t) - \mu)(X((t^\lambda - \tau\lambda)^{1/\lambda}) - \mu)] \\ &= E[(X((s^\lambda + \tau\lambda)^{1/\lambda}) - \mu)(X(s) - \mu)] \\ &= B_X(\tau; \lambda). \end{aligned}$$

**Definition 2.3** Let  $\varepsilon(u)$  be a white noise process, i.e.,  $E[\varepsilon(u)] = 0$  and

$$E[\varepsilon(u)\varepsilon(u + \tau)] = C\delta(\tau),$$

where  $C$  is a positive constant and  $\delta$  is the dirac delta function. Also let  $a(t) = \varepsilon\left(\frac{t^\lambda - 1}{\lambda}\right)$  for  $t > 0$ . Then we shall refer to  $a(t)$  as “ $G(\lambda)$ -white” noise.

**Theorem 2.2** For a  $G(\lambda)$ -white noise process,  $a(t)$ ,

$$B_a(\tau; \lambda) = E[a(t)a((t^\lambda + \tau\lambda)^{1/\lambda})] = C\delta(\tau).$$

Proof: The result is obvious.

**Definition 2.4** Let  $X(t)$  be a  $G(\lambda)$ -stationary process. Then the  $G(\lambda)$ -spectrum is

$$G_X(f; \lambda) = \int_{-\infty}^{\infty} e^{-2\pi if\tau} B_X(\tau; \lambda) d\tau.$$

Thus, the  $G(\lambda)$ -spectrum is the Fourier transformation of the  $G(\lambda)$ -autocovariance. If  $\lambda = 1$ , the  $G(\lambda)$ -spectrum is the usual spectrum of  $X(t)$ . If  $\lambda \rightarrow 0$ , from Property 2.1,  $B_X(\tau; 0) = R_X(e^\tau)$  and given  $h = e^\tau$ ,

$$\begin{aligned} G_X(f; 0) &= \int_{-\infty}^{\infty} e^{-2\pi if\tau} B_X(\tau; \lambda) d\tau \\ &= \int_{-\infty}^{\infty} e^{-2\pi if\tau} R_X(e^\tau) d\tau \\ &= \int_0^{\infty} h^{-2\pi if} R_X(h) d\ln h \\ &= \int_0^{\infty} h^{-2\pi if - 1} R_X(h) dh. \end{aligned}$$

then  $G_X(f; \lambda)$  is the Mellin transform of  $R_X(h)$ , the M-autocovariance of  $X(t)$ . So the  $G(\lambda)$ -spectrum with  $\lambda = 0$  is the M-spectrum of  $X(t)$ . The  $G(\lambda)$ -spectral density is defined by

$$M_X(f; \lambda) = \frac{G_X(f; \lambda)}{B_X(0; \lambda)} = \frac{G_X(f; \lambda)}{\text{var}(X(t))}.$$

Actually, the  $G(\lambda)$ -spectrum of a  $G(\lambda)$ -stationary process,  $X(t)$ , is equal to the usual spectrum of the dual process  $Y(u)$ , i.e.,  $G_X(f; \lambda) = S_Y(f)$ .

### Example 2.1

Let  $t \in (0, \infty)$ ,  $\phi \sim \text{Uniform}(0, 2\pi)$ ,  $\lambda \in (-\infty, \infty)$ ,  $A$  and  $\beta$  be constant, and let  $a(t)$  be  $G(\lambda)$ -white noise. Then the process

$$X(t) = A \cos(2\pi\beta(\frac{t^\lambda - 1}{\lambda}) + \phi) + a(t),$$

is  $G(\lambda)$ -stationary. The dual is  $Y(u) = A \cos(2\pi\beta u + \phi) + \varepsilon(u)$ , which is of course well known to be stationary. Figure 1 illustrates the periodic behavior of this  $G(\lambda)$ -stationary process for different values of  $\lambda$ . The periods elongate over time when  $\lambda$  is 0 and 0.5; the process appears stationary in the usual sense with a fixed period when  $\lambda = 1$ ; the periods contract in time when  $\lambda$  is 2. In the following section, we define a measure of the variation of the period or frequency of a process.

## 3 Instantaneous Period and Instantaneous Frequency of a $G(\lambda)$ process

For non-stationary signals, whose spectral content vary with time, the frequency at a particular time will be described by the concept of instantaneous frequency (IF).

We consider first a simple sinusoidal signal of the form

$$X(t) = A\cos(\phi(t)), \quad (1)$$

where  $A$  is constant and  $\phi$  is the cumulative phase of the signal. A common definition (Boashash, 1992) of the instantaneous frequency(IF) is

$$f(t) = \frac{\phi'(t)}{2\pi}.$$

Although the IF shows the rate of change of  $\phi(t)$  per  $2\pi$  units, it does not specify how many cycles there are for a given time interval, which may be our real interest in practice. Moreover, given a signal, in order to get its IF, we have to find the analytic signal using a Hilbert transformation, which is difficult to calculate. For some processes, the definition of IF may not have physical interpretation (Boashash, 1992). Therefore, we give a new definition of IF and a definition of the instantaneous period for  $G(\lambda)$  processes.

**Definition 3.1** *Let  $f$  be a function such that  $f(t) = f(g^{-1}(g(t) + \tau))$  for any  $t$ , where  $g$  is any monotonic function. Then  $f$  is said to be a  $G$ -periodic function with the  $G$ -period  $\tau$ , and the general instantaneous period (GIP) of the function  $f$ , denoted  $\ell(t; g, \tau)$ , is defined by*

$$\ell(t; g, \tau) = |g^{-1}(g(t) + \tau) - t|.$$

**Definition 3.2** *The general instantaneous frequency (GIF) of a function  $g(t)$  is*

$$f(t; \phi, \tau) = \frac{1}{\ell(t; g, \tau)}.$$

Note that when  $g$  is a monotonically increasing function,

$$\ell(t; g, \tau) = g^{-1}(g(t) + \tau) - t,$$

and  $\ell(t; g, \tau)$  is the length of the cycle starting at  $t$ . When  $g$  is a monotonically decreasing function,

$$\ell(t; g, \tau) = t - g^{-1}(g(t) + \tau),$$

and  $\ell(t; g, \tau)$  is the length of the cycle ending at  $t$ . If  $\ell(t; g, \tau)$  is increasing over time then this indicates elongation of the cyclical behavior. For a  $G(\lambda)$ -stationary process such as the one given in Example 2.1 with a single periodic component, it is easy to show that the GIP of the  $G(\lambda)$ -stationary process is

$$\ell(t; g, \tau) = \begin{cases} (t^\lambda + \lambda\tau)^{\frac{1}{\lambda}} - t & \lambda \neq 0, \\ t(e^\tau - 1) & \lambda \rightarrow 0 \end{cases} \quad (2)$$

and the corresponding GIF is

$$f(t; g, \tau) = \begin{cases} 1/((t^\lambda + \lambda\tau)^{\frac{1}{\lambda}} - t) & \lambda \neq 0, \\ 1/t(e^\tau - 1) & \lambda \rightarrow 0, \end{cases} \quad (3)$$

where  $\tau$  is the  $G$ -period of the  $G(\lambda)$ -stationary process  $X(t)$ . Actually,  $\tau$  is also the period of the dual process of  $g(t)$ . We will refer to the GIP and GIF of the  $G(\lambda)$ -stationary processes using the notation  $\ell(t; \lambda, \tau)$  and  $f(t; \lambda, \tau)$  since  $\lambda$  is the only parameter of the function  $g$ . If  $\lambda = 1$ ,  $\ell(t; 1, \tau) = \tau$  and  $f(t; 1, \tau) = 1/\tau$ . The GIP is a constant value  $1/\tau$  for any  $t > 0$ . This is consistent with the standard result that a weakly stationary process has fixed periods. When  $\lambda = 0$ , then  $e^\tau$  and  $1/\tau$  become the  $M$ -period and  $M$ -frequency respectively defined by Gray, Vijverberg and Woodward (2004). The GIP for  $\lambda = 0$  is linear in time.

As already described in Section 2, different values of  $\lambda$  result in different periodic behavior of the process. Figure 2 describes how the GIP changes over time for different  $\lambda$ 's. If  $0 < \lambda < 1$ , the GIP is a monotonically increasing concave function, while for  $\lambda < 0$ , it is monotonically increasing and convex. If  $\lambda > 1$ , the GIP is monotonically

decreasing and convex. Since GIF is defined as the reciprocal of GIP, the pattern for GIF is clear. Therefore, the class of  $G(\lambda)$ -stationary processes is quite flexible and can represent processes with a wide range of time-varying frequency behavior.

Although the definitions of GIF and IF are clearly not the same, there is an important relationship.

**Theorem 3.1** *When  $\phi$  is a monotonically continuous function, IF is the first order of the Taylor series approximation to the GIF. For  $G(\lambda)$ -stationary processes,  $IF = GIF$  if  $\lambda = 1$ , and  $IF \rightarrow GIF$  as  $t \rightarrow \infty$  if  $\lambda \neq 1$ .*

Given a sinusoidal signal of the form  $X(t) = A\cos(\phi(t))$  where  $A$  is constant, the IF is

$$f_1(t) = \frac{1}{2\pi}\phi'(t),$$

and the GIP is given by

$$l_2(t; \phi, 2\pi) = \phi^{-1}(\phi(t) + 2\pi) - t, \quad (4)$$

where we assume  $\phi(t)$  is a monotonically increasing function on  $t$ . The corresponding GIF is

$$f_2(t; \phi, 2\pi) = \frac{1}{l_2(t; \phi, 2\pi)}.$$

Using Taylor expansion of  $\phi^{-1}(y + 2\pi)$  about  $y$ , we obtain

$$\phi^{-1}(y + 2\pi) = \phi^{-1}(y) + \sum_{n=1}^{\infty} (2\pi)^n \frac{(\phi^{-1})^{(n)}(y)}{n!}.$$

Letting  $y = \phi(t)$ , then we have  $(\phi^{-1})'(y) = (\phi^{-1})'(\phi(t)) = 1/\phi'(t)$ ,  $(\phi^{-1})^{(2)}(y) = (\phi^{-1})^{(2)}(\phi(t)) = -\frac{\phi^{(2)}(t)}{(\phi'(t))^3}$  and  $(\phi^{-1})^{(3)}(y) = (\phi^{-1})^{(3)}(\phi(t)) = \frac{3\phi^{(2)}(t) - \phi^{(3)}(t)\phi'(t)}{(\phi'(t))^5}$ . Therefore,

$$\phi^{-1}(\phi(t) + 2\pi) = t + 2\pi \frac{1}{\phi'(t)} - \frac{(2\pi)^2}{2!} \frac{\phi^{(2)}(t)}{(\phi'(t))^3} + \frac{(2\pi)^3}{3!} \frac{3\phi^{(2)}(t) - \phi^{(3)}(t)\phi'(t)}{(\phi'(t))^5} + \dots$$

Then from Equation (4)

$$l_2(t; \phi, 2\pi) = 2\pi \frac{1}{\phi'(t)} - \frac{(2\pi)^2}{2!} \frac{\phi^{(2)}(t)}{(\phi'(t))^3} + \frac{(2\pi)^3}{3!} \frac{3\phi^{(2)}(t) - \phi^{(3)}(t)\phi'(t)}{(\phi'(t))^5} + \dots$$

If we use the first order Taylor expansion, then we obtain  $l_2(t; \phi, 2\pi) \approx 2\pi/\phi'(t)$  and  $f_2(t; \phi, 2\pi) = 1/l_2(t; \phi, 2\pi) \approx \phi'(t)/2\pi = f_1(t)$ . If  $\phi(t)$  is a linear function of  $t$ , i.e.,  $\phi(t) = at + b$ , then  $\phi^{(n)}(t) = 0$  for  $n \geq 2$ . We have  $l_2(t; \phi, 2\pi) = 2\pi/\phi'(t)$ , and  $f_2(t; \phi, 2\pi) = 1/l_2(t; \phi, 2\pi) = \frac{\phi'(t)}{2\pi} = f_1(t)$ . For  $\phi(t) = \frac{t^\lambda - 1}{\lambda}$ , when  $\lambda \neq 1$ , then  $l_2(t; \phi, 2\pi) \rightarrow \frac{2\pi}{\phi'(t)}$  as  $t \rightarrow \infty$  and  $f_2(t; \phi, 2\pi) = \frac{1}{l_2(t; \phi, 2\pi)} \rightarrow f_1(t)$  as  $t \rightarrow \infty$ .

For deterministic signals such as the simple sinusoidal signal mentioned previously, the GIP can be obtained empirically by peak detection. Figure 3(a) shows data from the model

$$X(t) = 10\cos(6\pi\ln(t)). \quad (5)$$

In Figure 3(a),  $t_1$  is the location of a peak, and  $a_1$  is the length between two consecutive peaks starting at  $t_1$ . So,  $a_1$  is the GIP at  $t_1$ . Also,  $a_2$  is the GIP at  $t_2$ . Figure 3(b) shows the plot of locations of peaks *vs* the corresponding GIP. This graph is a straight line which we expect due to the form of the GIP when  $\lambda = 0$ . However, for multicomponent signals, the notion of a single-valued GIP becomes meaningless, and a break-down into its components is needed. As with the definition of the spectrum of a stationary process, the GIP of a stochastic process is defined based on its autocovariance instead of the process itself (Jiang, 2003). If the periodic behavior in the process is very clear, the peak detection is still a valid method, and we also refer to results obtained in this manner as empirical general instantaneous periods (EGIP). Figure 3(c) shows a realization of the model

$$X(t) = A_1 \cos(2\pi\beta_1 \frac{(t + \Lambda)^\lambda - 1}{\lambda}) + A_2 \cos(2\pi\beta_2 \frac{(t + \Lambda)^\lambda - 1}{\lambda}) + a(t), \quad (6)$$

where  $A_1 = 10$ ,  $A_2 = 5$ ,  $\beta_1 = 5$ ,  $\beta_2 = 4.5$ ,  $\lambda = 0.5$ ,  $\Lambda = 100$  and  $a(t) \sim N(0,1)$ , and Figure 3(d) shows its EGIP obtained by peak detection. Although the model describes a process with two frequency components, and we can still obtain the pattern of GIP based on the EGIP. For the above example, theoretically,  $\lambda = 0.5$ , and the GIP corresponding to each component should be a monotonic increasing concave function. The EGIP in Figure 3(d) follows this pattern very well.

## 4 $G(p, q; \lambda)$ Processes

A typical class of  $G(\lambda)$ -stationary processes is  $G(p, q; \lambda)$  processes.

**Definition 4.1** *The process  $X(t)$  defined by*

$$\prod_{i=1}^p (t^{1-\lambda} D - \alpha_i) X(t) = \prod_{j=1}^q (t^{1-\lambda} D - \beta_j) a(t), \quad (7)$$

where  $t > 0$ ,  $D$  is the differential operator,  $\alpha_i$  and  $\beta_j$  are constants,  $\lambda \in (-\infty, \infty)$ ,  $p, q = 1, 2, 3, \dots$ , and  $a(t)$  is  $G(\lambda)$ -white noise, is referred to as the continuous  $G(p, q; \lambda)$  process. The equation

$$\prod_{i=1}^p (r - \alpha_i) = 0$$

is referred to as the characteristic equation of  $X(t)$ .

Remark: We will take the derivative here to be in the mean square sense. Equation (7) has a unique  $G(\lambda)$ -stationary solution and more correctly we define  $X(t)$  to be that solution. The  $G(\lambda)$ -stationarity or stability condition for the continuous  $G(p, q, \lambda)$  processes are that  $p > q \geq 0$  and that the real parts of the  $\alpha_i$ 's are negative. For a  $G(\lambda)$ -stationary process,  $X(t)$ , the time domain must be greater than 0, i.e.  $t > 0$ .

When  $\lambda = 1$ , a  $G(p, q; \lambda)$  process,  $X(t)$ , is actually a subset of a regular continuous ARMA( $p, q$ ) process with  $t > 0$ . In practice, we only observe and analyze processes at  $t > 0$ . We then call  $X(t)$  in the case the “observable” continuous ARMA( $p, q$ ) process. When  $\lambda = 0$ ,  $X(t)$  is the continuous Euler( $p, q$ ) process defined by Choi, Gray and Woodward (2003). The  $G(\lambda)$ -spectrum of  $X(t)$  is

$$G_X(f; \lambda) = \sigma_a^2 \frac{\prod_{j=1}^q |(\beta_j - i2\pi f)|^2}{\prod_{k=1}^p |(\alpha_k - i2\pi f)|^2}, \quad -\infty < f < \infty.$$

The stationary dual to  $X(t)$  is given by  $Y(u) = X(t)$ , where

$$u = \begin{cases} \ln(t) & \text{if } \lambda = 0, \\ \frac{t^\lambda - 1}{\lambda} & \text{otherwise.} \end{cases}$$

and  $Y(u)$  satisfies

$$\prod_{i=1}^p (D - \alpha_i)Y(u) = \prod_{j=1}^q (D - \beta_j)\varepsilon(u), \quad u \in (-\infty, \infty),$$

i.e.,  $Y(u)$  is a continuous ARMA( $p, q$ ) process.

When  $q = 0$ , we define Equation (7) be  $\prod_{i=1}^p (t^{1-\lambda}D - \alpha_i)X(t) = a(t)$ . Then  $G(p, 0; \lambda)$  processes are referred to  $G(p; \lambda)$  processes. If  $\lambda = 1$ , the  $G(p; \lambda)$  process can be considered to be the “observable” classical continuous AR( $p$ ) process; if  $\lambda = 0$ , it is the continuous  $p$ th-order Euler process (Gray and Zhang, 1988). Since the dual is given by  $Y(u) = X(t)$ , where  $u = \frac{t^\lambda - 1}{\lambda}$ , we have

$$\prod_{i=1}^p (D - \alpha_i)Y(u) = \varepsilon(u).$$

Also, we have following theorem:

**Theorem 4.1** *If  $X(t)$  is a continuous  $G(p; \lambda)$  process such as*

$$\prod_{i=1}^p (t^{1-\lambda}D - \alpha_i)X(t) = a(t), \quad t > 0,$$

then the  $G(\lambda)$ -autocovariance of  $X(t)$ , denoted by  $B_X(\tau; \lambda)$ , satisfies the differential equation

$$\prod_{i=1}^p (D - \alpha_i) B_X(\tau; \lambda) = 0.$$

$X(t)$  is  $G(\lambda)$ -stationary if the real parts of the  $\alpha_i$ 's are less than zero.

Proof: The result follows at once from Theorem 2.1 and the standard result.

## 5 Discretizing the Continuous $G(\lambda)$ -Stationary Process

Physical phenomena that are continuous stochastic processes are frequently observed at discrete time periods. From the observed data, inference is made regarding the underlying continuous process. The problem of discrete sampling from the usual continuous stationary process has been studied by Bartlett (1946) and Phadke and Wu (1974). They both showed that discretization of a continuous ARMA( $p, q$ ) process at an equally spaced sampling intervals results in a discrete ARMA( $p, p - 1$ ) process. For the  $G(\lambda)$ -stationary process, similar results can be obtained.

**Theorem 5.1** *If the data set  $X(t_k)$  is sampled from a  $G(\lambda)$ -stationary process  $X(t)$  at the time points  $t_k$ , i.e.,  $t_k = ((k + \zeta)\Delta\lambda + 1)^{1/\lambda}$ , where  $k=0,1,2,\dots$ , with  $\Delta > 0$  and  $\zeta > -\frac{1}{\Delta\lambda}$ , then  $Z_k = X(t_k)$  is a discrete stationary process. Also,  $C_Z(h) = B_X(h\Delta; \lambda)$ , where  $C_Z$  and  $B_X$  are the usual autocovariance of the process  $Z_k$  and  $G(\lambda)$ -autocovariance of the continuous process  $X(t)$ , respectively.  $Z_k$  will be referred to as the discrete dual of  $X(t)$  at the sample rate  $\Delta$ .*

Proof: Since  $Z_k = X(t_k)$ , then  $E(Z_k) = E(X(t_k)) = \mu$ ,  $\text{var}(Z_k) = \text{var}(X(t_k)) = \sigma^2$  and  $C_Z(h) = E[(Z_k - \mu)(Z_{k+h} - \mu)] = E[(X(t_k) - \mu)(X(t_{k+h}) - \mu)]$ . Letting  $s = t_k =$

$((k + \zeta)\Delta\lambda + 1)^{1/\lambda}$ , we have  $k = \frac{s^\lambda - 1}{\Delta\lambda} - \zeta$  and

$$\begin{aligned} t_{k+h} &= ((k + h + \zeta)\Delta\lambda + 1)^{1/\lambda} \\ &= (h\Delta\lambda + s^\lambda)^{1/\lambda}. \end{aligned}$$

Then, from Definition 2.1, we can obtain  $C_Z(h) = E[(X(t_k) - \mu)(X(t_{k+h}) - \mu)] = E[(X(s) - \mu)(X((h\Delta\lambda + s^\lambda)^{1/\lambda}) - \mu)] = B_X(h\Delta; \lambda)$ , which only depends on  $h$  given  $\Delta$ . Therefore,  $Z_k$  is a stationary process.

**Corollary 5.1** *If  $Y(u)$  is the continuous dual of  $X(t)$ , then  $C_Z(k) = C_Y(k\Delta)$ .*

Remark:  $t_k$  here is referred to as  $G(\lambda)$ -time scale. Since

$$k = \frac{t_k^\lambda - 1}{\Delta\lambda} - \zeta,$$

we also refer to it as the discrete Box-Cox transformation for  $t_k$ . Physically,  $t_0 = (\zeta\Delta\lambda + 1)^{1/\lambda}$  is the time of the first sampled data value. When  $\lambda = 1$ , then  $t_k = k\Delta + (\zeta\Delta + 1)$ , and thus sampling is at equally spaced time intervals  $\Delta$ . When  $\lambda = 0$ , then  $X(t)$  is a continuous Euler( $p, q$ ) process, and  $t_k = \lim_{\lambda \rightarrow 0} ((k + \zeta)\Delta\lambda + 1)^{1/\lambda} = e^{(k+\zeta)\Delta} = e^{\Delta\zeta}(e^\Delta)^k$ . Letting  $h = e^\Delta$  and  $A = e^{\Delta\zeta}$ , then the  $G(\lambda)$  time scale is

$$t_k = Ah^k, \tag{8}$$

where  $A > 0$ . Note that when  $\zeta$  is an integer, the  $G(\lambda)$ -time scale for  $\lambda = 0$  is the Euler time scale, and  $A = e^{\Delta\zeta}$  and  $h = e^\Delta$  are referred to as the offset and the sampling rate of the discrete M-stationary process, respectively (Gray, Vijverberg and Woodward, 2004).

For a  $G(\lambda)$ -stationary process  $X(t)$  and its stationary dual  $Y(u)$ , the discrete dual  $Z_k$  of  $X(t)$  can be also considered to be the equally spaced sample obtained from  $Y(u)$  at interval  $\Delta$ .

When the data are sampled from a continuous stationary process at an equally spaced interval  $\Delta$ , the frequency  $f_N = \frac{1}{2\Delta}$  is called the Nyquist frequency and is the highest frequency that can be detected by the sample. The Nyquist frequency of the discrete dual corresponds to a  $G(\lambda)$ -Nyquist frequency for the  $G(\lambda)$ -stationary process.

**Definition 5.1** *If the data are sampled from the  $G(\lambda)$  stationary process  $X(t)$  at  $t_k = ((k + \zeta)\Delta\lambda + 1)^{1/\lambda}$ , then  $f_N(\Delta; \lambda) = 1/2\Delta$  is called the  $G(\lambda)$ -Nyquist frequency.*

When  $\lambda = 1$ , the  $G(\lambda)$ -Nyquist frequency is the usual Nyquist frequency. When  $\lambda = 0$ , the corresponding  $G(\lambda)$ -Nyquist frequency is  $f_N(\Delta; 0) = 1/2\Delta = 1/(2\ln h)$ , where  $h$  is the sample rate for the Euler process (Gray, Vijverberg and Wooward, 2004).

**Theorem 5.2** *When the  $G(\lambda)$ -Nyquist frequency is greater than the highest frequencies corresponding to the characteristic equation of a  $G(p, q; \lambda)$  process,  $X(t)$ , where  $0 \leq q < p$ , the discrete dual of  $X(t)$  at a  $G(\lambda)$ -sampling interval  $\Delta$  is a discrete ARMA( $p, p - 1$ ) process. The coefficients of the discrete ARMA( $p, p - 1$ ) model depend both on the  $G(p, q; \lambda)$  model and the  $G(\lambda)$ -sampling interval,  $\Delta$ .*

Proof: The result follows at once from Theorem 2.1 and the result in Phadke and Wu (1974).

Remark: Given a  $G(p, q; \lambda)$  process, the coefficients of the discrete dual process depend on the  $G(\lambda)$ -sampling interval,  $\Delta$ . The discrete dual process is a discrete ARMA( $p, r$ ) process, where  $0 \leq r \leq p - 1$ .

### Example 5.1

Consider the continuous stationary process

$$(t^{1-\lambda}D - \alpha_1)(t^{1-\lambda}D - \alpha_2)X(t) = (t^{1-\lambda}D - \beta)a(t), \quad t > 0,$$

where  $\text{var}(a(t)) = \sigma_a^2$ . If we sample data from  $X(t)$  at time points  $t_k = ((k + \zeta)\Delta\lambda + 1)^{1/\lambda}$ , then the corresponding discrete ARMA(2,1) model is given by

$$X_k - \phi_1 X_{k-1} - \phi_2 X_{k-2} = \varepsilon_k - \theta \varepsilon_{k-1},$$

where  $\phi_1 = e^{\alpha_1 \Delta} + e^{\alpha_2 \Delta}$ , and  $\phi_2 = -e^{(\alpha_1 + \alpha_2) \Delta}$ , and  $\theta$  is defined as

$$\begin{cases} \theta = 0 & \text{if } \phi_1 \gamma_0 + (\phi_2 - 1) \gamma_1 = 0 \\ \theta^2 - (\phi_1 + \frac{\gamma_0 - \phi_1 \gamma_1 - \phi_2 \gamma_2}{\phi_1 \gamma_0 + (\phi_2 - 1) \gamma_1}) \theta + 1 = 0, |\theta| < 1 & \text{otherwise.} \end{cases}$$

The variance of  $\varepsilon_t$  is then given by

$$\sigma_\varepsilon^2 = \begin{cases} \frac{\phi_1 \gamma_0 + (\phi_2 - 1) \gamma_1}{\theta} & \text{if } \theta \neq 0, \\ \gamma_0 - \phi_1 \gamma_1 - \phi_2 \gamma_2 & \text{if } \theta = 0, \end{cases}$$

and

$$\gamma_k = \gamma(k\Delta) = \frac{\alpha_1^2 - \beta^2}{2\alpha_1(\alpha_2^2 - \alpha_1^2)} e^{\alpha_1 |k| \Delta} + \frac{\beta^2 - \alpha_2^2}{2\alpha_2(\alpha_2^2 - \alpha_1^2)} e^{\alpha_2 |k| \Delta}.$$

According to Theorem 5.2, if the continuous model is  $G(2; \lambda)$ , the corresponding discrete stationary dual is still a discrete ARMA(2,1) process. The coefficients of the MA term of the continuous processes only affect the white noise variance and the value of the MA term in the discrete stationary dual. Note that here we have

$$\begin{aligned} \phi_1^2 + 4\phi_2 &= (e^{\alpha_1 \Delta} + e^{\alpha_2 \Delta})^2 - 4e^{(\alpha_1 + \alpha_2) \Delta} \\ &= (e^{\alpha_1 \Delta} - e^{\alpha_2 \Delta})^2. \end{aligned}$$

So, if  $\alpha_1$  and  $\alpha_2$  are complex conjugates, then  $\phi_1^2 + 4\phi_2 = (e^{\alpha_1\Delta} - e^{\alpha_2\Delta})^2 < 0$ , in Region I of Figure 4, and the characteristic function of the discrete process has two complex roots. If  $\alpha_1$  and  $\alpha_2$  are both negative real number (since the continuous process is stationary, real parts of  $\alpha_1$  and  $\alpha_2$  must be negative), then  $\phi_1^2 + 4\phi_2 = (e^{\alpha_1\Delta} - e^{\alpha_2\Delta})^2 \geq 0$ , in Region II of Figure 4, and the characteristic function of the discrete process has two positive real roots.

The following example shows how to obtain the corresponding continuous ARMA(2,1) model, i.e., the G(2,1;1) process, of the discrete ARMA(2,1) model.

### Example 5.2

Given the stationary discrete model

$$X_k - \phi_1 X_{k-1} - \phi_2 X_{k-2} = \varepsilon_k - \theta \varepsilon_{k-1}, \quad (9)$$

the autocovariance function is

$$\gamma_k = C_1 G_1^{|k|} + C_2 G_2^{|k|}, \quad (10)$$

where  $G_1$  and  $G_2$  are given by

$$\frac{\phi_1}{2} \pm \frac{\sqrt{\phi_1^2 + 4\phi_2}}{2},$$

$C_1$  and  $C_2$  are given by  $C_1 = \frac{\gamma_1 - \gamma_0 G_2}{G_1 - G_2}$  and  $C_2 = \gamma_0 - C_1$ , and

$$\gamma_0 = \frac{(\phi_2 - 1)(\theta^2 + 1) + 2\phi_1\theta}{(\phi_2 + 1)(\phi_1^2 - (\phi_2 - 1)^2)} \sigma_\varepsilon^2$$

and

$$\gamma_1 = \frac{(\phi_1 - \theta)(\phi_1\theta - 1) - \phi_2^2\theta}{(\phi_2 + 1)(\phi_1^2 - (\phi_2 - 1)^2)} \sigma_\varepsilon^2.$$

Assuming the discrete process is obtained by sampling the continuous process,  $X(t)$ , at equally spaced interval  $\Delta$  and  $\gamma(u)$  is the autocorrelation function of  $X(t)$ , then we have

$$\gamma(k\Delta) = \gamma_k = C_1 G_1^{|k|} + C_2 G_2^{|k|}.$$

Under the assumption that the continuous process and its discrete sample have the same autocorrelation structure, for any lag  $u$ , the autocorrelation function,  $\gamma(u)$ , of  $X(t)$  can be written as

$$\gamma(u) = C_1 e^{\frac{\ln(G_1)}{\Delta}|u|} + C_2 e^{\frac{\ln(G_2)}{\Delta}|u|}. \quad (11)$$

Letting  $\alpha_1 = \frac{\ln(G_1)}{\Delta}$  and  $\alpha_2 = \frac{\ln(G_2)}{\Delta}$ , then we get

$$\gamma(u) = C_1 e^{\alpha_1|u|} + C_2 e^{\alpha_2|u|}. \quad (12)$$

For different values of  $\phi_1$  and  $\phi_2$ , we have the following four cases:

(i)  $\phi_1^2 + 4\phi_2 < 0$ :  $\phi_1$  and  $\phi_2$  lie in Region I of Figure 4.

In this case,  $G_1$  and  $G_2$  are complex conjugates, i.e.,

$$G_1 = a - bi,$$

$$G_2 = a + bi.$$

$G_1$  and  $G_2$  can be expressed in polar form as  $G_1 = R(\cos\omega + i\sin\omega) = Re^{i\omega}$  and  $G_2 = Re^{-i\omega}$ , where  $\tan\omega = b/a$  and  $R = \sqrt{a^2 + b^2}$ . In our case,  $R = \sqrt{a^2 + b^2} = \sqrt{G_1 G_2} = \sqrt{-\phi_2}$  and  $\cos\omega = \frac{a}{\sqrt{a^2 + b^2}} = \frac{\phi_1}{2\sqrt{-\phi_2}}$ . Then we have  $\ln G_1 = \ln R + i\omega$ ,  $\ln G_2 = \ln R - i\omega$  and

$$\alpha_1 = \frac{\ln R + i\omega}{\Delta},$$

$$\alpha_2 = \frac{\ln R - i\omega}{\Delta}.$$

Substituting into Equation (12), we get

$$\gamma(u) = C_1 e^{\frac{\ln_{R+i\omega}|u|}{\Delta}} + C_2 e^{\frac{\ln_{R-i\omega}|u|}{\Delta}}.$$

From the standard result, this is the autocorrelation function of the continuous ARMA(2,1) process, i.e.,

$$(D - \alpha_1)(D - \alpha_2)X(t) = (D - \beta)a(t), \quad (13)$$

where  $a(t)$  is continuous white noise. Since the discrete process (9) is stationary,  $R = \sqrt{-\phi_2} < 1$ . So, the real parts of  $\alpha_1$  and  $\alpha_2$  are negative, and the continuous process (13) is stationary. The spectrum  $G(f)$  of the process (13) is known to be

$$G(f) = \frac{|i2\pi f - \beta|^2}{|(i2\pi f - \alpha_1)(i2\pi f - \alpha_2)|^2} \sigma_a^2, \quad (14)$$

where  $\sigma_a^2$  is the variance of  $a(t)$ . Since  $G(f)$  is also the Fourier transform of  $\gamma(u)$ , we have

$$\begin{aligned} G(f) &= \int_{-\infty}^{\infty} \gamma(u) e^{-i2\pi f u} du \\ &= C_1 \int_{-\infty}^{\infty} e^{\alpha_1|u|} e^{-i2\pi f u} du + C_2 \int_{-\infty}^{\infty} e^{\alpha_2|u|} e^{-i2\pi f u} du \\ &= C_1 \int_{-\infty}^{\infty} e^{\alpha_1|u| - i2\pi f u} du + C_2 \int_{-\infty}^{\infty} e^{\alpha_2|u| - i2\pi f u} du \\ &= \frac{-C_1}{\alpha_1 + i2\pi f} + \frac{-C_1}{\alpha_1 - i2\pi f} + \frac{-C_2}{\alpha_2 + i2\pi f} + \frac{-C_2}{\alpha_2 - i2\pi f} \\ &= -2(C_1\alpha_1 + C_2\alpha_2) \frac{|i2\pi f + \sqrt{\frac{C_1\alpha_1\alpha_2^2 + C_2\alpha_2\alpha_1^2}{C_1\alpha_1 + C_2\alpha_2}}|^2}{|(i2\pi f - \alpha_1)(i2\pi f - \alpha_2)|^2}. \end{aligned} \quad (15)$$

Comparing Equation (14) and Equation (15), we can get  $\beta = -\sqrt{\frac{C_1\alpha_1\alpha_2^2 + C_2\alpha_2\alpha_1^2}{C_1\alpha_1 + C_2\alpha_2}}$  and  $\sigma_a^2 = -2(C_1\alpha_1 + C_2\alpha_2)$ . Note that to have a one-to-one correspondence between model (13) and the spectrum (14), it is necessary that the real part of  $\beta$  be nonpositive.

(ii)  $\phi_1^2 + 4\phi_2 \geq 0$ ,  $\phi_2 \leq 0$  and  $\phi_1 \geq 0$ :  $\phi_1$  and  $\phi_2$  lie in Region II of Figure 4.

In this case, both  $G_1$  and  $G_2$  are positive real numbers, Equation (12) can be written as

$$\gamma(u) = C_1 e^{\frac{\ln(|G_1|)}{\Delta}|u|} + C_1 e^{\frac{\ln(|G_2|)}{\Delta}|u|}.$$

This is similar to case (i). The corresponding continuous process is ARMA(2, 1), i.e., model (13) with  $\alpha_1 = \frac{\ln(|G_1|)}{\Delta}|u|$ ,  $\alpha_2 = \frac{\ln(|G_2|)}{\Delta}|u|$ ,  $\beta = -\sqrt{\frac{C_1\alpha_1\alpha_2^2 + C_2\alpha_2\alpha_1^2}{C_1\alpha_1 + C_2\alpha_2}}$  and  $\sigma_a^2 = -2(C_1\alpha_1 + C_2\alpha_2)$ . Since both  $\alpha_1$  and  $\alpha_2$  are real numbers,  $X(t)$  here does not have periodic behavior.

(iii)  $\phi_1^2 + 4\phi_2 \geq 0$ ,  $\phi_2 > 0$ :  $\phi_1$  and  $\phi_2$  lie in Region III of Figure 4.

In this case,  $G_1$  and  $G_2$  are both real numbers. Since  $G_1G_2 = \phi_2 < 0$ , we let  $G_1 < 0$  and  $G_2 > 0$ . since  $\ln(G_1) = \ln(|G_1|) + i\pi$  or  $\ln(G_1) = \ln(|G_1|) - i\pi$ , Equation (12) can be written as

$$\begin{aligned} \gamma(u) &= C_1 e^{\frac{\ln(G_1)}{\Delta}|u|} + C_2 e^{\frac{\ln(G_2)}{\Delta}|u|} \\ &= \frac{C_1}{2} e^{\frac{\ln(|G_1|) - i\pi}{\Delta}|u|} + \frac{C_1}{2} e^{\frac{\ln(|G_1|) + i\pi}{\Delta}|u|} + C_2 e^{\frac{\ln(G_2)}{\Delta}|u|}. \end{aligned}$$

Letting  $\alpha_{11} = \frac{\ln(|G_1|) - i\pi}{\Delta}$ ,  $\alpha_{12} = \frac{\ln(|G_1|) + i\pi}{\Delta}$  and  $\alpha_2 = \frac{\ln(G_2)}{\Delta}$ , the corresponding continuous process is a continuous ARMA(3,2) process, i.e.,

$$\begin{aligned} (D - \alpha_{11})(D - \alpha_{12})(D - \alpha_2)X(t) &= (D - \beta_1)(D - \beta_2)a(t) \\ &= \frac{d^2a(t)}{dt^2} + b_0 \frac{da(t)}{dt} + b_1 a(t), \end{aligned} \quad (16)$$

where  $b_0 = -(\beta_1 + \beta_2)$  and  $b_1 = \beta_1\beta_2$ . The spectrum,  $G(f)$ , of the continuous process (16) is known to be

$$\begin{aligned} G(f) &= \frac{|(i2\pi f - \beta_1)(i2\pi f - \beta_2)|^2}{|(i2\pi f - \alpha_{11})(i2\pi f - \alpha_{12})(i2\pi f - \alpha_2)|^2} \sigma_a^2 \\ &= \frac{(2\pi f)^4 + (b_0^2 - 2b_1)(2\pi f)^2 + b_1^2}{|(i2\pi f - \alpha_{11})(i2\pi f - \alpha_{12})(i2\pi f - \alpha_2)|^2} \sigma_a^2 \end{aligned} \quad (17)$$

Since  $G(f)$  is also the Fourier transform of  $\gamma(u)$ , we have

$$\begin{aligned}
G(f) &= \int_{-\infty}^{\infty} \gamma(u) e^{-i2\pi f u} du \\
&= \frac{C_1}{2} \int_{-\infty}^{\infty} e^{\alpha_{11}|u|} e^{-i2\pi f u} du + \frac{C_1}{2} \int_{-\infty}^{\infty} e^{\alpha_{12}|u|} e^{-i2\pi f u} du + C_2 \int_{-\infty}^{\infty} e^{\alpha_2|u|} e^{-i2\pi f u} du \\
&= \frac{C_1}{2} \int_{-\infty}^{\infty} e^{\alpha_{11}|u| - i2\pi f u} du + \frac{C_1}{2} \int_{-\infty}^{\infty} e^{\alpha_{12}|u| - i2\pi f u} du + C_2 \int_{-\infty}^{\infty} e^{\alpha_2|u| - i2\pi f u} du \\
&= \frac{-\frac{C_1}{2}}{\alpha_{11} + i2\pi f} + \frac{-\frac{C_1}{2}}{\alpha_{11} - i2\pi f} + \frac{-\frac{C_1}{2}}{\alpha_{12} + i2\pi f} + \frac{-\frac{C_1}{2}}{\alpha_{12} - i2\pi f} \\
&\quad + \frac{-C_2}{\alpha_2 + i2\pi f} + \frac{-C_2}{\alpha_2 - i2\pi f} \\
&= -(C_1\alpha_{11} + C_1\alpha_{12} + 2C_2\alpha_2) \left[ \frac{(2\pi f)^4}{|(i2\pi f - \alpha_{11})(i2\pi f - \alpha_{12})(i2\pi f - \alpha_2)|^2} \right. \\
&\quad + \frac{\left( \frac{C_1\alpha_2^2(\alpha_{11} + \alpha_{12}) + C_1(\alpha_{11} + \alpha_{12})\alpha_{11}\alpha_{12} + 2C_2\alpha_2\alpha_{11}\alpha_{12}}{C_1\alpha_{11} + C_1\alpha_{12} + 2C_2\alpha_2} \right) (2\pi f)^2}{|(i2\pi f - \alpha_{11})(i2\pi f - \alpha_{12})(i2\pi f - \alpha_2)|^2} \\
&\quad \left. + \frac{C_1(\alpha_{11} + \alpha_{12})\alpha_{11}\alpha_{12}\alpha_2^2 + C_2\alpha_2\alpha_{11}^2\alpha_{12}^2}{C_1\alpha_{11} + C_1\alpha_{12} + 2C_2\alpha_2} \right]
\end{aligned}$$

Comparing it with (17), we can get  $\sigma_a^2 = -(C_1\alpha_{11} + C_1\alpha_{12} + 2C_2\alpha_2)$ ,

$$b_1 = \sqrt{\frac{C_1(\alpha_{11} + \alpha_{12})\alpha_{11}\alpha_{12}\alpha_2^2 + C_2\alpha_2\alpha_{11}^2\alpha_{12}^2}{C_1\alpha_{11} + C_1\alpha_{12} + 2C_2\alpha_2}}$$

and

$$b_0 = \sqrt{\frac{C_1\alpha_2^2(\alpha_{11} + \alpha_{12}) + C_1(\alpha_{11} + \alpha_{12})\alpha_{11}\alpha_{12} + 2C_2\alpha_2\alpha_{11}\alpha_{12}}{C_1\alpha_{11} + C_1\alpha_{12} + 2C_2\alpha_2}} + 2b_1.$$

Note that the frequency corresponding to the complex roots,  $\alpha_{11}$  and  $\alpha_{12}$ , of the continuous process,  $X(t)$ , is  $f_0 = \frac{1}{2\pi} \frac{\pi}{\Delta} = \frac{1}{2\Delta}$ , which is equal to the Nyquist frequency  $\frac{1}{2\Delta}$ . That is, if the Nyquist frequency is equal to highest frequency corresponding to the complex roots of the continuous process, sampling the continuous ARMA(3,2) then results in a discrete ARMA(2,1) process.

(iv)  $\phi_1^2 + 4\phi_2 \geq 0$ ,  $\phi_2 \leq 0$  and  $\phi_1 < 0$ :  $\phi_1$  and  $\phi_2$  lie in Region IV of Figure 4.

In this case, both  $G_1$  and  $G_2$  are negative, and Equation (12) can be written as

$$\begin{aligned}\gamma(u) &= \frac{C_1}{2} e^{\frac{\ln(|G_1|) - i\pi}{\Delta} |u|} + \frac{C_1}{2} e^{\frac{\ln(|G_1|) + i\pi}{\Delta} |u|} \\ &\quad + \frac{C_2}{2} e^{\frac{\ln(|G_2|) - i\pi}{\Delta} |u|} + \frac{C_2}{2} e^{\frac{\ln(|G_2|) + i\pi}{\Delta} |u|}.\end{aligned}$$

The corresponding continuous process then is a continuous ARMA(4,3) process (Jiang, 2003), and the two frequencies corresponding to the complex roots of the continuous process are both equal to the Nyquist frequency  $\frac{1}{2\Delta}$ .

From Example 5.2, if the  $G(\lambda)$ -Nyquist frequency is equal to the highest frequency corresponding to the complex roots of the characteristic equation of the  $G(p, q; \lambda)$  process, then we have following results.

**Theorem 5.3** *Let  $X(t)$  be a  $G(p, q; \lambda)$  process and let  $f_0$  be the highest frequency corresponding to the complex roots  $\alpha_i$ 's of characteristic equation of  $X(t)$ . Then sampling  $X(t)$  at  $G(\lambda)$  time scale with  $G(\lambda)$ -Nyquist frequency equal to  $f_0$  results in a discrete ARMA( $p - \frac{m}{2}; p - \frac{m}{2} - 1$ ) process, where  $m$  is the number of the complex roots  $\alpha_i$ .*

Proof: The result follows at once from Theorem 2.1 and we replace the continuous ARMA(2, 1) process in Example 5.2 by the continuous  $G(p, q; \lambda)$  process.

Remark: In this article, we only consider  $G(\lambda)$ -sampling interval  $\Delta$  such that  $f_0 \leq \frac{1}{2\Delta}$ , where  $f_0$  is the highest frequency corresponding to the complex roots the characteristic function of continuous processes. From a practical viewpoint, the assumption that  $f_0$  is not greater than the Nyquist frequency or  $G(\lambda)$ -Nyquist frequency in  $G(\lambda)$ -stationary processes is trivial, since a properly planned data collection would invariably have  $\Delta$  small enough so that  $f_0 \leq \frac{1}{2\Delta}$ . Under the assumption, for a given

$G(\lambda)$ -time scale and a given discrete ARMA( $p, q$ ) model, there is a unique continuous  $G(\lambda)$ -stationary process. This implies that for the continuous  $G(\lambda)$ -stationary process, modeling can be based on the discrete dual. The realization of the continuous  $G(p, q; \lambda)$  process can be approximated by the realization of its discrete dual process with a small  $G(\lambda)$ -sampling interval. Given a  $G(p, q; \lambda)$  process, different  $G(\lambda)$ -sample intervals results in different discrete dual process. The following theorem shows the relationship between the system frequency of the original  $G(p, q; \lambda)$  processes and that of the discrete dual processes.

**Theorem 5.4** *Suppose  $Z_{1k}$  and  $Z_{2k}$  are discrete duals of the continuous  $G(\lambda)$ -stationary process  $X(t)$  at sample rates  $\Delta_1$  and  $\Delta_2$ , respectively. If both  $G(\lambda)$ -Nyquist frequencies,  $f_{1N}(\Delta_1; \lambda) = 1/(2\Delta_1)$  and  $f_{2N}(\Delta_2; \lambda) = 1/(2\Delta_2)$ , are greater than the highest frequency corresponding to the complex  $\alpha_i$ 's of the process  $X(t)$ , then  $Z_{1k}$  and  $Z_{2k}$  have the same number of system frequencies, i.e.,  $f_{1i}$  and  $f_{2i}$ ,  $i = 1, 2, \dots, m$ , where  $m$  is the number of pairs of complex roots of characteristic function of  $X(t)$ . Moreover,  $\frac{f_{1i}}{\Delta_1} = \frac{f_{2i}}{\Delta_2} = \frac{b_i}{2\pi}$ , where  $b_i$  is the imaginary part of those complex roots.*

Proof: The result follows at once from Theorem 2.1 and the standard result for continuous ARMA ( $p, q$ ) processes (Priestley, 1981).

### Example 5.3

Consider the continuous  $G(2, 1; \lambda)$  process  $X(t)$  such that

$$(t^{1-\lambda}D - \alpha_1)(t^{1-\lambda}D - \alpha_2)X(t) = (t^{1-\lambda} - \beta)a(t),$$

where  $\alpha_1$  and  $\alpha_2$  are complex conjugates. From Example 5.1, the corresponding discrete ARMA(2,1) models for  $Z_{1k}$  and  $Z_{2k}$  are

$$Z_{1k} - \phi_{11}Z_{1(k-1)} - \phi_{12}Z_{1(k-2)} = \varepsilon_{1k} - \theta_{11}\varepsilon_{1(k-1)}$$

and

$$Z_{2k} - \phi_{21}Z_{2(k-1)} - \phi_{22}Z_{2(k-2)} = \varepsilon_{2k} - \theta_{21}\varepsilon_{2(k-1)},$$

where  $\phi_{11} = e^{\alpha_1\Delta_1} + e^{\alpha_2\Delta_1}$ ,  $\phi_{12} = -e^{(\alpha_1+\alpha_2)\Delta_1}$ ,  $\phi_{21} = e^{\alpha_1\Delta_2} + e^{\alpha_2\Delta_2}$  and  $\phi_{22} = -e^{(\alpha_1+\alpha_2)\Delta_1}$ . Then system frequencies associated with  $Z_{1k}$  and  $Z_{2k}$  are

$$f_{11} = \frac{1}{2\pi} \cos^{-1}\left(\frac{\phi_{11}}{2\sqrt{-\phi_{12}}}\right)$$

and

$$f_{21} = \frac{1}{2\pi} \cos^{-1}\left(\frac{\phi_{21}}{2\sqrt{-\phi_{22}}}\right).$$

Suppose  $\alpha_1 = a + ib$  and  $\alpha_2 = a - ib$ , where  $a$  and  $b$  are real numbers. Then we obtain

$$\frac{\phi_{11}}{2\sqrt{-\phi_{12}}} = \frac{e^{\alpha_1\Delta_1} + e^{\alpha_2\Delta_1}}{2\sqrt{e^{(\alpha_1+\alpha_2)\Delta_1}}} = \frac{e^{\alpha_1\Delta_1} + e^{\alpha_2\Delta_1}}{2e^{\frac{\alpha_1+\alpha_2}{2}\Delta_1}} = \frac{e^{a\Delta_1}(e^{ib\Delta_1} + e^{-ib\Delta_1})}{2e^{a\Delta_1}} = \cos(b\Delta_1),$$

$f_{11} = \frac{b\Delta_1}{2\pi}$  and  $f_{21} = \frac{b\Delta_2}{2\pi}$ . So we have  $\frac{f_{11}}{\Delta_1} = \frac{f_{21}}{\Delta_2} = \frac{b}{2\pi}$ .

Remark: The  $G(\lambda)$ -system frequencies of a  $G(p, q; \lambda)$  process are the system frequency of its continuous dual. When  $\lambda = 0$ , i.e.,  $X(t)$  is an Euler( $p, q$ ) process, then the system frequencies of the discrete dual divided by the  $G(\lambda)$ -sampling interval are referred to as M-system frequencies(Gray, Vijverberg and Woodward, 2004) of the discrete process. If the sampling is sufficiently fast, the M-system frequencies of the discrete dual process are independent of the sampling rate and are equal to the M-system frequency—the imaginary part of corresponding complex roots of the characteristic function, of the continuous M-stationary process divided by  $2\pi$ . Theorem 5.3 thus implies that the  $G(\lambda)$ -system frequency of the  $G(p, q; \lambda)$  can be obtained from its discrete dual.

## 6 Instantaneous Spectrum of $G(\lambda)$ -stationary Processes

The GIP and GIF of a stochastic process measure the periods and the frequencies at a given time. From Section 1, the  $G(\lambda)$ -autocorrelation of a  $G(\lambda)$ -stationary process is

$$\rho_X(\tau; \lambda) = E[(X(t) - \mu)(X((t^\lambda + \tau\lambda)^{1/\lambda}) - \mu)]/\text{var}(X(t)).$$

Let  $\rho_X^*(r, t)$  be the autocorrelation between  $X(t)$  and  $X(t+r)$ . That is,

$$\rho_X^*(r, t) = E[(X(t) - \mu)(X(t+r) - \mu)]/\text{var}(X(t)).$$

Letting  $\tau = \frac{(t+r)^\lambda - t^\lambda}{\lambda}$ , we have

$$\begin{aligned} \rho_X(\tau; \lambda) &= E[(X(t) - \mu)(X((t^\lambda + \tau\lambda)^{1/\lambda}) - \mu)]/\text{var}(X(t)) \\ &= E[(X(t) - \mu)(X(t+r) - \mu)]/\text{var}(X(t)) \\ &= \rho_X^*(r, t). \end{aligned}$$

That is,

$$\rho_X^*(r, t) = \rho_X(\tau; \lambda) = \rho_X\left(\frac{(t+r)^\lambda - t^\lambda}{\lambda}; \lambda\right), \quad (18)$$

which also means that the usual ACF of the  $G(\lambda)$ -stationary process depends not only on the lag  $r$  but also on the time  $t$ . For example, if  $\lambda = 0$ ,  $X(t)$  is an M-stationary process and

$$\begin{aligned} \lim_{\lambda \rightarrow 0} \rho_X^*(r, t) &= \lim_{\lambda \rightarrow 0} \rho_X\left(\frac{(t+r)^\lambda - t^\lambda}{\lambda}; \lambda\right) \\ &= \rho_X(\ln(1+r/t); 0). \end{aligned} \quad (19)$$

The GIP of  $G(\lambda)$ -stationary process,  $X(t)$ , at time  $t$  is defined as the length of the first cycle of  $\rho_X^*(r, t)$ .

Figure 5(a) shows a realization from a  $G(2,1;0)$  model with the discrete dual

$$X_k - 1.732X_{k-1} + 0.98X_{k-2} = \varepsilon_k, \quad (20)$$

with sample rate  $h = 1.0055$  and  $\text{var}(\varepsilon_k) = 1$ . Note that we only show the part of the realization from  $t = 101$  to  $t = 300$ . Figure 5(b) shows the ACF,  $\rho_X^*(r, t)$ , calculated by Equation (19), of  $X(t)$  at time  $t = 101$ . Then  $a1$  is the GIP of  $X(t)$  at time  $t = 101$ , and  $f = 1/a1$  is the GIF of  $X(t)$  at time  $t = 101$ . However, if the process contains several periodic components, it is not easy to observe the GIP and GIF from the ACF of the process. Figure 6(a) shows a realization from a  $G(4, 3; 0)$  model with discrete dual

$$(1 - 1.4B + 0.98B^2)(1 - 1.732B + 0.98B^2)X_k = \varepsilon_k - 0.6\varepsilon_{k-1}, \quad (21)$$

at sample rate  $h = 1.0055$  and  $\text{var}(\varepsilon_k) = 1$ . Similar to Figure 5(a), we also just show the part of the realization from  $t = 101$  to  $t = 300$ . Figure 6(b) shows its ACF,  $\rho_X^*(r, t)$ , calculated using Equation (18), of  $X(t)$  at time  $t = 101$ . As we expect, it is not easy to observe GIF from the ACF of the process directly. However, there are two clear peaks shown in the  $G(\lambda)$ -spectrum(Figure 6(c)) of the  $G(4, 3; 0)$  model. This indicates that  $X(t)$  has two dominant frequencies, which change over time.

Figure 7(a) shows a realization from the  $G(6, 5; 0)$  model with discrete dual

$$(1 - 1.4385B + 0.9944B^2)(1 - 0.0054B + 0.9697B^2)(1 + 1.3006B + 0.9226B^2)X_k = \varepsilon_k, \quad (22)$$

at sample rate  $h = 1.0055$  and  $\text{var}(\varepsilon_k) = 1$ . Figure 7(b) shows the ACF. It seems that its cyclical behavior is similar to what we saw in Figure 5(b). However, its  $G(\lambda)$ -spectrum(Figure 7(c)) clearly shows that the data have three dominant frequencies, which change over time, while the data in Figure 5(a) only have one dominant fre-

quency. Therefore, the  $G(\lambda)$ -spectrum can provide information about frequencies that are clearly unable to see from other methods.

As mentioned in Section 2, the  $G(\lambda)$ -spectrum of a  $G(\lambda)$ -stationary process equals to the usual spectrum of its stationary dual. For example, Figures 8(a)-(b) show the realization and  $G(\lambda)$ -spectrum of a  $G(2,1;0)$  process with discrete stationary dual

$$X_k - 1.4X_{k-1} + 0.98X_{k-2} = \varepsilon_k, \quad (23)$$

at sample rate  $h = 1.0055$  and innovation variance 1, respectively. Figures 8(c)-(d) show a realization and the  $G(\lambda)$ -spectrum from a  $G(2,1;2)$  model with the discrete dual in Equation (23) at the  $G(\lambda)$ -sampling interval  $\Delta = 200.5$  and innovation variance 1, respectively. From Figure 8(a), the data are elongating, while they are compacting in Figure 8(c). However, their  $G(\lambda)$ -spectra look similar. Thus, the  $G(\lambda)$ -spectrum only provides the information about frequencies of the data series at the transformed time scale, but cannot tell how those frequencies change over time. Therefore, we define the instantaneous spectrum for the  $G(\lambda)$ -stationary process.

**Definition 6.1** *If  $X(t)$  is a  $G(\lambda)$ -stationary process and  $G_X(f^*; \lambda)$  is its  $G(\lambda)$ -spectrum, then the instantaneous spectrum of  $X(t)$  at time  $t$  is*

$$\mathcal{S}(f, t; \lambda) = \begin{cases} G_X(0; \lambda) & \text{if } f = 0, \\ G_X\left(\frac{1}{\ln(\frac{1}{tf} + 1)}; \lambda\right) & \text{if } f \neq 0 \text{ and } \lambda = 0, \\ G_X\left(\frac{\lambda}{(t+1/f)^\lambda - t^\lambda}; \lambda\right) & \text{otherwise.} \end{cases}$$

The instantaneous spectrum actually measures the spectrum of the  $G(\lambda)$ -stationary process at a specific time  $t$ , and it is based on its  $G(\lambda)$ -spectrum. In practice, we use the spectrum of the discrete dual to approximate the  $G(\lambda)$ -spectrum of the underlying continuous process. If  $X$  is an equally spaced realization from the continuous

$G(\lambda)$ -stationary process  $X_k, k = 1, 2, \dots, n$ , then we have  $t = \Lambda + k$  and  $f \leq 0.5$ , where  $\Lambda$  is the offset of the data.

Figure 9 shows a realization of a  $G(2,1;0)$  process which has discrete dual process (23) with sample rate  $h = 1.0055$  and innovation variance 1. Figure 10 shows the corresponding instantaneous spectrum, where the x-axis represents time, and the y-axis is the frequency. The gray scale represents the power at the corresponding frequency and time. The darker the color, the higher the power. For example, the location of the most powerful frequency of the data changes from about 0.22 at time  $t = 101$  to 0.09 at time  $t = 250$ . This indicates that at time  $t = 101$ , the length of the next cycle is  $l = 1/0.22 \approx 4$ , and the length of next cycle at time  $t = 250$  is  $l = 1/0.09 \approx 11$ . If we count those lengths in the data series in Figure 11, we will obtain these results. Figure 11 shows a realization from a  $G(2, 1; 1)$  model. In its instantaneous spectrum plot(Figure 12), the dark line stays at about  $f = 0.125$  over the time period from  $t = 101$  to  $t = 300$ . That is, the  $G(\lambda)$ -stationary process with  $\lambda = 1$ , i.e., the traditional stationary process, has constant frequency or frequencies, which is consistent with the common sense of the frequency of the stationary processes. Actually, the usual spectrum based on the Fourier transformation could be treated as a special case of the instantaneous spectrum at  $\lambda = 1$ . Figure 13 and Figure 14 show a realization and the corresponding instantaneous spectrum of a  $G(2, 1; 2)$  model. The data are obviously compacting, and the dominant frequency monotonically increases over time. Figure 15 shows a realization from a  $G(6,5;0)$  process. From its instantaneous spectrum (Figure 16), the data set has three dominant frequencies, which monotonically decrease over time. From the data set, it seems that the GIP at  $t = 101$  is about 4 and the dominant GIP at  $t = 250$  is about 11. This is reflected by the most dominant frequency, where  $f = 0.22$  at  $t = 101$  and  $f = 0.09$  at

$t = 250$ . Figure 17 and Figure 18 show a realization and the corresponding instantaneous spectrum from  $G(6,5;1)$  process, i.e., ARMA(6,5) process, respectively, where both three dominant frequencies keep constant over time. Figure 19 and Figure 20 show a case for  $\lambda = 2$ , for which the data are compacting and all of the dominant frequencies increase over time. In addition to displaying the change of the frequency, the instantaneous spectrum provides information about “aliasing”. In Figure 16, the top line of the three dominant frequencies starts at about  $t = 135$ , and before that time, the data only shows two frequencies. This is due to the fact that one of the instantaneous frequencies(top line) is so high(larger than 0.5) before  $t = 135$  that the discrete data sampled at the fixed interval cannot detect it when  $t < 135$ . It follows that in sampling continuous time series, adequate care must be taken to ensure that a high enough sampling frequency is chosen so that aliasing will be avoided.

The instantaneous spectrum relates to time-frequency analysis, which has been the object of intense research activity in the last decade. The common methods such as the Gabor transforms and wavelet transforms provide some good results for the time-frequency data. The instantaneous spectrum analysis provides another approach under the assumption that the data are  $G(\lambda)$ -stationary. We will not discuss the relationship between these methods here, but it will be a topic of further research.

## 7 The Offset Problem and Equally Spaced Realizations

The correlation function of a stationary process is, of course, not a function of time, while the correlation function of a  $G(\lambda)$ -stationary process does change with time when  $\lambda \neq 1$ . The impact of this, from a data point of view, is that in order to

properly model the process, one needs some estimate of the location of the initial observation. To be more specific, suppose for a time shift,  $\Lambda > 0$ , and  $t > 0$  that  $X_1(t) = X(t - \Lambda)$  is a  $G(\lambda)$ -stationary process with  $\lambda = 0$ , i.e.,  $X_1(t)$  is an M-stationary process. The process  $X(t)$  is not M-stationary but is referred to as the shifted M-stationary process (Gray and Zhang, 1988). The time shift,  $\Lambda$ , will be referred to as the origin offset of the process  $X(t)$ . The stationary dual of  $X_1(t)$  is  $Y_1(u)$ , i.e.,  $Y_1(u) = X_1(e^u)$ . However, the process  $Y(u) = X(e^u) = X_1(e^u + \Lambda)$  here is definitely not stationary. That is, given the process  $X(t)$ ,  $t > 0$ , we cannot obtain the stationary dual through a logarithmic time transformation without adjusting for the origin offset. Actually, we can regard  $X(t)$  as the observed subset of  $X_1(t)$ . Thus, it is reasonable to use the stationary dual,  $Y_1(u)$ , of  $X_1(t)$  as the stationary dual of  $X(t)$ . In order to obtain the stationary dual of  $X_1(t)$ , we then use the time transformation function  $\ln(t + \Lambda)$  instead of  $u = \ln(t)$ , i.e.,  $Y(u) = X(e^u - \Lambda) = X_1(e^u) = Y_1(u)$ . For the  $G(\lambda)$ -stationary process we use the time transformation function,

$$u = \frac{(t + \Lambda)^\lambda - 1}{\lambda}, \Lambda > 0,$$

instead of  $u = \frac{t^\lambda - 1}{\lambda}$ . Note that if  $\lambda = 1$ , the  $G(\lambda)$ -stationary process is the observable stationary process, and the correlation function does not depend on  $t$ . Therefore, we do not need to estimate the origin offset,  $\Lambda$ .

## 7.1 Generating Realizations from a $G(p, q; \lambda)$ Model

For most time series modeling, observations are usually discrete and taken at equally spaced time intervals, i.e., hourly, daily, monthly and so on. In order to imitate the practice, we need to generate realizations of the  $G(\lambda)$ -stationary process at equally spaced time intervals. Suppose the equally spaced time interval is 1. The discrete

realization is  $X_k = X(\Lambda + k)$ , where  $k = 1, 2, 3, \dots$ , where  $X(t)$  is the continuous process and  $\Lambda$  is the offset. For processes such as

$$X(t) = A\cos(2\pi\beta(\frac{t^\lambda - 1}{\lambda}) + \phi) + a(t),$$

this is easy. However, generating equally spaced data from a  $G(p, q; \lambda)$  process is not as straightforward. The main problem is that we cannot generate a continuous realization from a  $G(p, q; \lambda)$  process. When  $\lambda = 1$ , the  $G(p, q; \lambda)$  process is an observable continuous stationary ARMA( $p, q$ ) process. From Theorem 5.2, its equally spaced sample at a sampling interval  $\Delta$  is the discrete ARMA( $p, p - 1$ ). Therefore, an equally spaced realization of the continuous ARMA( $p, q$ ) process can be generated from the corresponding discrete ARMA( $p, p - 1$ ). For the continuous  $G(p, q; \lambda)$  process, given the  $G(\lambda)$ -time scale,  $t_k = ((k + \zeta)\Delta\lambda + 1)^{1/\lambda}$ ,  $k = 0, 1, 2, \dots, n$ , we first generate the realization,  $Z_k$ , from its corresponding discrete dual. Then we obtain an approximation of the continuous realization of  $X(t)$  based on

$$\hat{X}(t) = \begin{cases} Z_k & \text{if } t = t_k \\ Z(t; \{Z_k\}) & \text{otherwise} \end{cases}, \quad (24)$$

where  $Z$  is the interpolation function. One way to do such interpolation is to use standard smoothing procedures. If we use linear interpolation,

$$Z(t; \{Z_k\}) = \frac{t_{k+1} - t}{t_{k+1} - t_k} Z_k + \frac{t - t_k}{t_{k+1} - t_k} Z_{k+1},$$

where  $k$  is the largest integer such that  $t_k < t$ . Then the finally equally spaced sample can be obtained by  $X_k = \hat{X}(\Lambda + k)$ . In (24), we use interpolation to approximate the continuous process, so there will be interpolation errors in the final equally spaced sample. However, when the  $G(\lambda)$ -sampling interval  $\Delta$  is very small, the interpolation error is minor. Our experience showed that the effect of interpolation methods is minor.

The following procedure gives the details for generating an equally spaced discrete realization with interval 1 and sample size  $n$  from a continuous  $G(p, q; \lambda)$  process with the offset,  $\Lambda$ .

1. Setup a  $G(\lambda)$ -time scale,  $t_k = ((k + \zeta)\Delta\lambda + 1)^{1/\lambda}$ , where  $\Delta$  is the  $G(\lambda)$ -sampling interval. Usually, we let  $\Delta = \frac{(\Lambda+n)^\lambda - (\Lambda+1)^\lambda}{m(n-1)\lambda}$ ,  $\zeta = \frac{(\Lambda+1)^\lambda - 1}{\Delta\lambda}$  and  $k = 0, 1, 2, \dots, m(n-1)$ . So we have  $t_0 = \Lambda + 1$  and  $t_{m(n-1)} = \Lambda + n$ .
2. Find the corresponding discrete dual at the  $G(\lambda)$ -sampling interval  $\Delta$  as we did in Example 5.1.
3. Generate the realization,  $Z_k$ ,  $k = 0, 1, 2, \dots, m(n-1)$ , from the discrete dual.
4. Obtain the approximation  $\hat{X}(t)$  for  $\Lambda + 1 \leq t \leq \Lambda + n$  using Equation (24)
5. Obtain the equally spaced realization using  $X_k = \hat{X}(\Lambda + k)$ ,  $k = 1, 2, \dots, n$ .

Note that  $m$  here is referred to as the generating rate which determines how many data values are in the realization of the discrete dual to approximate the continuous process. The higher  $m$  is, the lower the approximation error in step (4) will be. Thus, large  $m$  results in small interpolation error. Typically, we set  $m = 20$  for which the interpolation error is minor.

### Example 7.1

Let  $X(t)$  be a continuous Euler(2,1) process given by

$$(tD - \alpha_1)(tD - \alpha_2)X(t) = (tD - \beta)a(t), \quad (25)$$

for  $t > 0$ , where  $\alpha_1 = -1.844 - 92.31i$ ,  $\alpha_2 = -1.844 + 92.31i$ ,  $\beta = 433.5$  and  $\text{var}(a(t)) = 34.4$ . We generated a discrete realization at equally spaced intervals of length 1 and sample

size  $n = 200$  from  $X(t)$  with the offset  $\Lambda = 100$ . The  $G(\lambda)$ -time scale in this case is

$$\begin{aligned} t_k &= \lim_{\lambda \rightarrow 0} ((k + \zeta)\Delta\lambda + 1)^{1/\lambda} \\ &= Ah^k, \end{aligned}$$

where  $A = e^{\Delta\zeta}$  and  $h = e^\Delta$ . From step 1,  $\Delta = \lim_{\lambda \rightarrow 0} \frac{(\Lambda+n)^\lambda - (\Lambda+1)^\lambda}{m(n-1)\lambda} = \frac{\ln(\Lambda+n) - \ln(\Lambda+1)}{m(n-1)}$  and  $\zeta = \lim_{\lambda \rightarrow 0} \frac{(\Lambda+1)^\lambda - 1}{\Delta\lambda} = \frac{\ln(\Lambda+1)}{\Delta}$ . When  $m = 20$ , obtain  $A = e^{\Delta\zeta} = e^{\ln(\Lambda+1)} = 101$  and  $h = e^\Delta = \left(\frac{\Lambda+n}{\Lambda+1}\right)^{\frac{1}{m(n-1)}} = \left(\frac{300}{101}\right)^{\frac{1}{20 \cdot 199}} = 1.000274$ . Using the procedure described in Example 5.1, the discrete dual of  $X(t)$  given  $h = 1.000274$  is

$$Z_k - 1.99835Z_{k-1} + 0.99899Z_{k-2} = 0.88796\varepsilon_k, \quad (26)$$

where  $\text{var}(\varepsilon_k) = 0.01058$ . A discrete realization (see Figure 21(a)) with  $m(n-1) = 3981$  data points then is generated from the discrete dual process, i.e., model (26). An approximation of the continuous realization of  $X(t)$  for  $101 \leq t \leq 300$  can be obtained by using Equation (24) (see Figure 21(b)). Figure 21(c) shows the equally spaced data set with length 200 using the equation  $X_k = \hat{X}(\Lambda + k)$ ,  $k = 1, 2, \dots, 200$ .

## 7.2 Modeling Equally Spaced Data

We assume that the underlying  $G(\lambda)$ -stationary process is continuous. Since most observed data sets are equally spaced, it is necessary to interpolate to obtain the discrete dual. Given the equally spaced discrete sample  $X_t$ ,  $t = 1, 2, 3, \dots, N$ , and the offset of the process,  $\Lambda$ , the interpolated continuous sample is

$$\hat{X}(t + \Lambda) = \begin{cases} X_t & \text{if } t = 1, 2, 3, \dots, N, \\ X(t; \{X_t\}) & \text{otherwise,} \end{cases}$$

where  $X$  is the interpolation function. If we use linear interpolation, then

$$X(t; \{X_t\}) = ([t] + 1 - t)X_{[t]} + (t - [t])X_{[t]+1},$$

where  $[t]$  represents the integer part of  $t$ . There are many interpolation methods that could be used such as local regression estimation (Loess) and spline estimation. Different interpolation methods may yield different results. However, when the cyclic behavior is clearly visible in the original discrete data, the effect of the interpolation methods on the final model is minor. The discrete dual of  $\hat{X}(t)$ ,  $Z_k$ , is

$$Z_k = \hat{X}(((k + \zeta)\Delta\lambda + 1)^{1/\lambda} - \Lambda), k = 0, 1, 2, \dots$$

Usually, we let  $\Delta = \frac{(\Lambda+n)^\lambda - (\Lambda+1)^\lambda}{(m-1)\lambda}$  and  $\zeta = \frac{(\Lambda+1)^\lambda - 1}{\Delta\lambda}$ , where  $n$  is the sample size of the original equally spaced data, and  $m$  is the length of the discrete dual. In practice, we can fix either the sample rate  $\Delta$  or the number of the dual data values,  $m = n$ .

For example, given an equally spaced sample from an M-stationary process, i.e.,  $\lambda = 0$ , with offset  $\Lambda$ , we have

$$\begin{aligned} \Delta &= \lim_{\lambda \rightarrow 0} \frac{(\Lambda + n)^\lambda - (\Lambda + 1)^\lambda}{(m - 1)\lambda} \\ &= \frac{\ln(\Lambda + n) - \ln(\Lambda + 1)}{m - 1} \\ &= 1/(m - 1)\ln\left(\frac{\Lambda + n}{\Lambda + 1}\right) \end{aligned}$$

and  $\zeta = \lim_{\lambda \rightarrow 0} \frac{(\Lambda+1)^\lambda - 1}{\Delta\lambda} = \frac{\ln(\Lambda+1)}{\Delta}$ . The G( $\lambda$ )-time scale is  $t_k = Ah^k$ , where  $h = e^\Delta = \left(\frac{\Lambda+n}{\Lambda+1}\right)^{1/(m-1)}$ ,  $A = e^{\Delta\zeta} = \Lambda + 1$  and  $k = 0, 1, 2, \dots, m - 1$ . Often, we let  $m = n$ , i.e., we obtain  $n$  dual data values from the  $n$  equally spaced data values.

### Example 7.2

Figure 22(a) shows an equally spaced realization from model (25) with the length  $n = 200$ . The offset is  $\Lambda = 100$ . Figure 22(b) shows the discrete dual using linear interpolation with sampling rate  $h = 1.005486$ . Using Burg's algorithm to fit the

discrete dual, we obtain a discrete AR(2) model given by

$$Z_k - \hat{\mu} - 1.7448(Z_{k-1} - \hat{\mu}) + 0.9939(Z_{k-2} - \hat{\mu}) = a_k.$$

The estimated variance of  $a_k$  is 0.95, and  $\hat{\mu} = -0.164$ . The corresponding estimated Euler(2, 1) model is

$$(tD - \hat{\alpha}_1)(tD - \hat{\alpha}_2)(X(t) - \hat{\mu}) = (tD - \hat{\beta})\varepsilon(t), \quad (27)$$

for  $t > 0$ , where  $\hat{\alpha}_1 = -0.56 - i92.34$ ,  $\hat{\alpha}_2 = -0.56 + i92.34$ ,  $\hat{\beta} = -434.2$ , and  $\text{var}(\varepsilon(t)) = 32.3$ . Comparing these results with the parameters of the true model, (25), the estimated coefficients are close to the true values. In the continuous process such as the continuous ARMA process, the imaginary parts of the complex roots of the characteristic equation determine the frequencies of the data, and the real parts determine the energy of those frequencies. The closer the real part is to 0, the stronger the corresponding frequency. Therefore, model (27) does a very good job of estimating the cyclical behavior and the white noise variance of the true model (25). In Figure 22(c), the solid line is the M-spectrum of true model, while the dashed line is the M-spectrum based on model (27). Both have a peak at frequency, 14.7. Based on the spectrum, the estimated model fits the true model very well. Figure 22(d) shows the M-autocorrelation function of the true model, i.e., the solid line, and that of the estimated model. The periodic behavior of the M-autocorrelation function of the estimated model matches that of true model very well. Since the absolute value of the real parts of the complex roots of the true model are larger than that of the estimated model, the true M-spectrum has a higher peak than the estimated one.

### 7.3 Forecasting

There are several methods of forecasting after fitting a model to an equally spaced data set. The method used here is to forecast the discrete dual and then interpolate to obtain forecasts for  $X(t)$  at the equally spaced time points. Given an equally spaced data set  $X_t$ , suppose we know  $\Lambda$  and  $\lambda$ . In Section 7.2, we describe how to fit a  $G(\lambda)$  model such as a  $G(p, q; \lambda)$  model to the data. Then, for a  $G(\lambda)$ -time scale such as  $t_k = ((k + \zeta)\Delta\lambda + 1)^{1/\lambda}$ , we then can obtain its discrete dual to which a model such an  $ARMA(p, q)$  can fit. The procedure for obtaining the  $l$ -step ahead forecasts follows.

1. Calculate the corresponding maximum forecasting length,  $L$ , for the discrete dual data using equation

$$L = \left[ \frac{(\Lambda + n + l)^\lambda - 1}{\Delta} - m - 2 \right],$$

where  $\Lambda$  is the offset,  $n$  is the length of the original equally spaced data set,  $m$  is the length of the discrete dual,  $\Delta$  is the sampling interval,  $l$  is the maximum lag to forecast in the original series, and  $[ \ ]$  means the integer part.

2. Using the discrete  $ARMA(p, q)$  model and the discrete dual data,  $Z_0, Z_1, \dots, Z_{m-1}$ , obtain  $L$ -step ahead forecasts,  $Z_m, Z_{m+2}, \dots, Z_{m+L-1}$ .

3. Obtain the approximation,  $\hat{X}(t)$ , for  $t_{m-1} \leq t \leq t_{m+L-1}$  using equation

$$\hat{X}(t) = \begin{cases} Z_M & \text{if } t = t_M, \\ Z(t; \{Z_M\}) & \text{otherwise,} \end{cases}$$

where  $M = m - 1, m, \dots, m + L - 1$  and  $Z$  is the interpolation function.

4. Obtain the equally spaced  $l$ -step ahead forecasts by  $X_k = \hat{X}(\Lambda + k)$ ,  $k = n + 1, n + 2, \dots, n + l$ .

Figure 23 shows the  $l$ -step ahead forecasts using the model fit to the discrete dual data indicated by a dashed line with sign “+”, and using the usual AR model fitting the original equally spaced data indicated by a dotted line, respectively, where the solid line is the true value. Clearly, the  $G(\lambda)$  method produces better forecast performance than the typically AR model. In order to compare the forecast performance, we define

$$\text{Improvement} = \frac{\text{MSE of AR Model} - \text{MSE of } G(\lambda) \text{ Model}}{\text{MSE of AR Model}} \times 100\%,$$

where MSE indicates the mean squared-errors of the forecasts. The value of “Improvement” measures how much the  $G(\lambda)$  model reduces the forecast error. The maximum value of improvement is 100%, which means there is no forecast error using the  $G(\lambda)$  model. The value can be negative, which indicates that the forecast error increases when using the  $G(\lambda)$  model. The closer the value of improvement is to 100%, the better the forecast performance of the  $G(\lambda)$  model is compared to the AR model. In Table 1, we can see that the  $G(\lambda)$  model improves the forecast performance for Example 7.2 dramatically. This example demonstrates that for the time series which have frequencies monotonically changing over time, the  $G(\lambda)$  model can provide substantially better forecast than the usual ARMA model.

## 8 Estimation of $\lambda$ and $\Lambda$

We have shown that given equally spaced data and the values of  $\Lambda$  and  $\lambda$ , we can find the discrete dual,  $Z_k$ , at the sampling rate  $\Delta$  by interpolation. From Theorems 5.2 and 5.3, modeling  $Z_k$  yields the corresponding continuous model as shown in Example 5.2. Now we discuss the estimation of  $\lambda$  and  $\Lambda$ .

Table 1: Forecast Performance(MSE): AR model vs.  $G(\lambda)$  Model for Example 7.2

lags	AR Model	$G(\lambda)$ Model	Improvement
5	44.5	11.0	75%
10	52.6	22.5	57%
15	103.4	24.9	76%
20	150.2	29.4	80%
25	250.0	49.2	80%
30	309.5	55.9	82%
35	311.9	69.7	78%
40	360.3	111.4	69%
45	441.0	137.4	69%
50	434.2	148.6	66%
55	464.1	188.3	59%
60	502.4	207.8	59%

Gray and Vijverberg (2002) introduced a model-based method for estimating  $\Lambda$  for a shifted M-stationary process, which is the  $G(\lambda)$ -stationary process with  $\lambda = 0$ . They obtained an initial “guess” for the possible values of  $\Lambda$ , then found the value of  $\Lambda$  that minimized the sum of squared residuals (SSE) after fitting an AR model to the dual data. The method was based on finding the value of  $\Lambda$  that provides the model with minimum weighted SSE fit to the dual data. These results showed good forecast performance when roots of the characteristic function were close to unit circle. However, their estimator of  $\Lambda$  was dependent on the initial “guess”, and the resulting estimate was not always close to the true value in the simulation studies.

In order to narrow down the range of the origin and avoid error due to the initial “guess”, Choi (2003) investigated an alternative approach that balanced the number of dual data values in the first and last cycle. The approach works well when there is one frequency in the dual process and when this frequency can be clearly seen in the realization. In the case of multiple frequencies, it will be difficult to identify the first cycle and last cycle. Choi (2003) also used AIC instead of SSE as the criterion to estimate  $\Lambda$  in order to improve the comparison among models of different orders. From the simulation results, AIC not only tends to give models with fewer parameters, but also yields a more accurate and stable estimator of  $\Lambda$ . Both AIC and SSE are model-based. Different models or different modeling methods such as “Burg”, “Yule-Walker”, and “MLE” often result in very different estimators of  $\Lambda$ , especially when the roots of the characteristic equation are not very close to the unit circle. For  $G(\lambda)$ -stationary processes, we need to estimate both  $\lambda$  and  $\Lambda$ . Before any estimation, a reasonable step is to test  $H_0 : \lambda = 1$ . If  $\lambda = 1$ , the traditional stationary analysis is appropriate. Otherwise, we need to estimate the value of  $\lambda$  and  $\Lambda$ . The new approach for estimation discussed in this section is model-free and is based on the empirical general instantaneous period (EGIP) and the sample autocorrelation function(SACF) of the discrete dual data.

### 8.1 Testing $H_0 : \lambda = 1$ using SACF

Given a discrete time series  $X_t$ , the sample autocovariance function is

$$\hat{\gamma}_k = \frac{1}{n} \sum_{t=1}^{n-|k|} (X_t - \bar{X})(X_{t+|k|} - \bar{X}), k = 0, \pm 1, \pm 2, \dots, \pm(n-1),$$

and the sample autocorrelation function is  $\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$ . When  $\lambda = 1$ ,  $X_t$  is a discrete stationary process, and  $\hat{\gamma}_k$  and  $\hat{\rho}_k$  are estimators of its autocovariance and auto-

correlation function, respectively. When  $\lambda \neq 1$ ,  $\hat{\gamma}_k$  and  $\hat{\rho}_k$  have no direct physical interpretation. If we divide a time series realization from a  $G(\lambda)$ -stationary process into two sections, and then calculate the SACFs of the two parts,  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$ , respectively, the two SACFs can help test  $H_0 : \lambda = 1$  and identify  $\lambda$  and  $\Lambda$ . Figure 24(a) and (c) show 400 equally spaced data values generated from models:

$$(D - \alpha_1)(D - \alpha_2)X(t) = (D - \beta)a(t) \quad (28)$$

and

$$(t^{1/2}D - \alpha_1)(t^{1/2}D - \alpha_2)X(t) = (t^{1/2}D - \beta)a(t), \quad (29)$$

where  $\alpha_1 = -0.0101 - 0.506i$ ,  $\alpha_2 = -0.0101 + 0.506i$ ,  $\beta = 2.375$  and  $\text{var}(a(t)) = 0.188$ , respectively. When  $\lambda = 1$ , i.e. model (28), the data are stationary, and  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$  are very close to each other since both are the estimators of the autocorrelation of the same stationary process (Figure 24(b)). When  $\lambda = 0.5$ , i.e. model (29), the cyclical behavior of  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$  can be clearly different as seen in Figure 24(d). Actually, the two SACFs,  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$ , can be interpreted as the local autocorrelations of the first half series and the second half series respectively. For  $G(\lambda)$ -stationary processes, the theoretical local sample autocorrelation of the first half series is equal to that of the second half series if and only if  $\lambda = 1$ , i.e., the process is stationary. In Figures 24(c)-(d), when  $\lambda = 0.5$ ,  $\hat{\rho}_{1k}$  seems to have higher frequency than  $\hat{\rho}_{2k}$ . Therefore, given a time series, one approach to test  $H_0 : \lambda = 1$  is to test whether there is a significant difference between  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$ . Here we first use a graphical display to assess whether  $\lambda = 1$ . Given a time series  $X_k$ , the procedure follows:

1. Separate the data  $X_k$  into two parts, the first half,  $X_{1k}$ , and the second half,  $X_{2k}$ , and obtain the SACFs,  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$ , respectively.

2. Fit an autoregressive model to  $X_{1k}$ . Suppose the sample size of the full data set,  $X_k$ , is  $n$ . Generate  $R$  realizations of length  $n$  using the autoregressive model obtained for  $X_{1k}$ .
3. For each realization, calculate  $\hat{\rho}_{1k}$ , and obtain an empirical distribution of the  $\hat{\rho}_{1k}$  at each  $k$ .
4. For  $\hat{\rho}_{1k}$  at fixed lag,  $k$ , obtain a  $100(1-\alpha)\%$  probability interval for the  $\hat{\rho}_{1k}$  using the quantiles of the empirical distribution.
5. Using the  $100(1-\alpha/2)$ th and  $(100\alpha/2)$ th quantiles at each lag, an  $100\alpha$  envelope is obtained for  $\hat{\rho}_{1k}$  by connecting these quantiles.
6. Plot the envelope and  $\hat{\rho}_{2k}$ .

Here the value  $R$  is the number of bootstrap replications, and  $R = 199$  in the example. If  $X_k$  is from a stationary model with constant frequency, then  $\hat{\rho}_{2k}$  should fall largely within the envelope. This type of envelope was used by Tsay (1992) and Ripley (1977), and they will be referred to here as acceptance envelopes. The probability that the envelope contains the entire SACF is not  $1 - \alpha$  under the assumption that the data are from a stationary model. Thus the acceptance envelope considered is not a joint  $100(1-\alpha)\%$  confidence interval of the SACF but is mainly used as a guide. Figure 25(a) shows  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$  for a realization generated by model (28) and an acceptance envelope constructed by using  $\alpha=0.10$ , where the solid lines are the 5th percentile and 95th percentile of SACF, the dotted line is  $\hat{\rho}_{1k}$  and the dashed line with label “2” is  $\hat{\rho}_{2k}$ . Here the envelope contains  $\hat{\rho}_{2k}$  very well. However, from Figure 25(b), the stationarity of the realization from model (29) is questionable. Therefore, the bootstrap-based envelope provides a graphical display that can be used to assess

whether the data are from a model with constant period or frequency, i.e,  $\lambda = 1$ . As mentioned above, this graphical method is not to test  $H_0: \lambda = 1$  but simply provides a diagnostic display.

To test the null hypothesis  $H_0 : \lambda = 1$ , we measure the difference between  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$  by  $z_k = \hat{\rho}_{1k} - \hat{\rho}_{2k}$ ,  $i = 1, 2, \dots, K$ , where  $K$  is the maximum lag of SACFs. In the implementation here, we use  $K$  equal to one fourth of the total sample size. Under the null hypothesis:  $H_0: \lambda = 1$ , we have

$$\underline{\mathbf{Z}} \sim \text{normal}(\mathbf{0}, \underline{\mathbf{\Sigma}}),$$

where  $\underline{\mathbf{Z}} = (z_1, z_2, \dots, z_K)$ ,  $\mathbf{0}$  is a  $K \times 1$  vector with 0 as elements, and  $\underline{\mathbf{\Sigma}}$  is the covariance matrix

$$\underline{\mathbf{\Sigma}} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1K} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{K1} & \gamma_{K2} & \cdots & \gamma_{KK} \end{pmatrix},$$

where  $\gamma_{ij} = 2\text{cov}(\hat{\rho}_{1i}, \hat{\rho}_{1j})$ . It follows that  $\underline{\mathbf{Z}}' \underline{\mathbf{\Sigma}}^{-1} \underline{\mathbf{Z}} \sim \chi^2(K)$  under the null.

To test the difference between  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$  via parametric bootstraps, we propose a new statistic

$$Q = \sum_{k=0}^K (\hat{\rho}_{1k} - \hat{\rho}_{2k})^2.$$

For a stochastic process, typically,  $\underline{\mathbf{\Sigma}} \neq \mathbf{I}$ , and the statistic  $Q$  does not have a simple known distribution such as Chi-square or Gamma. The bootstrap method is used to generate the empirical distribution of  $Q$  under the assumption that the time series is stationary. The procedure follows:

1. Separate the time series  $X_k$  into two parts, the first half  $X_{1k}$  and the second half  $X_{2k}$ , and obtain the SACFs and the value of the  $Q$ -statistic.

2. Fit an autoregressive model to  $X_k$ , the entire series.
3. Generate  $R$  realizations using the model obtained in step (2) and calculate the value of the  $Q$ -statistic for each realization ( $R = 199$  in the example).
4. Obtain  $100(1-\alpha)$ th empirical quantiles,  $Q_{1-\alpha}$ , of  $Q$ .

If  $Q > Q_{1-\alpha}$ , we then reject the null hypothesis. Figures 25(c) and (d) give the resulting empirical distributions of  $Q$ 's for model (28) and model (29) and the estimated  $Q_{.95}$ 's (vertically dashed line). The observed value of  $Q$  is also shown with the arrow. In Figure 25(c) we show the case in which  $\lambda = 1$ , and in Figure 25(d), we show the case for  $\lambda = 0.5$ . It is clear that  $Q < Q_{.95}$  when  $\lambda = 1$ , and  $Q > Q_{.95}$ , indicating rejection of  $H_0 : \lambda = 1$ , when  $\lambda = 0.5$ .

## 8.2 Initial Estimation of $\Lambda$ and $\lambda$ using EGIP

After testing the null hypothesis, we need to either fit a stationary model to the data if we do not reject it or find the proper Box-Cox transformation for the time scale if we reject the null. Theoretically, given a  $G(\lambda)$ -stationary process, the dual is stationary if and only if we transform the time scale by the true  $\lambda$  and the true offset  $\Lambda$ . That is, the general instantaneous period(GINP) of the dual is a positive constant if and only if the true  $\lambda$  and  $\Lambda$  are identified. When data have clear periodic behavior, the empirical general instantaneous period(EGINP) estimated by peak detection can be used to estimate the GINP. The procedure we recommend involves fitting a linear regression line for the EGINP on the dual time scale. Thus, we will be searching for values of  $\Lambda$  and  $\lambda$  that are associated with a regression line with slope near zero.

The algorithm is as follows:

1. For the original equally spaced data, find the locations of peaks, i.e.,  $t_1, t_2, \dots, t_m$  by inspection.
2. For a given value of  $\lambda$ , let

$$u_i = \begin{cases} t_i & \text{if } \lambda = 1 \\ \ln(t_i + \Lambda) & \text{if } \lambda = 0 \\ \frac{(t_i + \Lambda)^\lambda - 1}{\lambda} & \text{otherwise} \end{cases}$$

and  $p_i = u_{i+1} - u_i, i = 1, 2, 3, \dots, m - 1$ .

3. Initially, letting  $\lambda = 1$ , find the corresponding  $m - 1$  pairs of  $p_i$  and  $u_i$  values, where  $p_i$  is also referred to as the EINP of the data at time  $u_i$ . Fit a simple linear regression model:  $p_i = bu_i + a$ , and test whether the slope  $b$  is significantly different from 0 at the significance level 95%.
4. Choose a value of  $\lambda$  from a selected range. If  $b > 0$  and the general trend of the EINP appears to be a straight line, we usually look at  $\lambda$ 's in the range  $(0, 1)$ , or perhaps even  $(-1, 1)$ . We consider  $\Lambda$ 's in the range  $(0, 200)$ , at first, and extend the range later if necessary. We would usually cover the selected range with about 11-21 values of  $\lambda$  and 20-50 values of  $\Lambda$ . We can divide up a portion of the interval more finely later. If  $b < 0$ , we would select the range for  $\lambda$  to be  $(1, 3)$ , or perhaps even  $(1, 5)$ .
5. For each combination of chosen  $\lambda$  and  $\Lambda$  values, calculate  $m - 1$  pairs  $u_i$  and  $p_i$ , fit a simple linear regression model:  $p_i = bu_i + a$  and record  $|b(\lambda, \Lambda)|$ , the absolute slope for the regression, and  $|t(b, \lambda, \Lambda)|$ , the absolute  $t$  value associated with the estimated slope.

6. Find all combinations of  $\lambda$ 's and  $\Lambda$ 's for which the slope,  $b$ , is not significantly different from 0.

The initial estimation will narrow the range of  $\lambda$  and  $\Lambda$  significantly, especially, when the periodic behavior is very clear in the data.

### 8.3 Estimating $\lambda$ and $\Lambda$ Using $Q$

As previously mentioned, for the  $G(\lambda)$ -stationary processes with offset  $\Lambda$ , the theoretical autocorrelation of the first half of the dual data is equal to that of the second half dual data if and only if the discrete dual is obtained using the true  $\lambda$  and  $\Lambda$ . Since the  $Q$ -statistic can be used to measure the difference between two SACFs, we use a numerical method to find the value of  $\lambda$  and  $\Lambda$  that minimize  $Q$ . The procedure follows.

1. Separate the time series  $X_k$  into two parts, the first half  $X_{1k}$  and the second half  $X_{2k}$ , and obtain the two corresponding SACFs and the value of the  $Q$ -statistic.
2. Choose the range of the values of  $\lambda$  and  $\Lambda$  from the initial estimation based on EGIP. Current practice is to cover the selected range with about 11-21 values of  $\lambda$  and 20-50 values of  $\Lambda$ . The interval can be divided up a more finely if needed.
3. Find the discrete dual for each combination of  $\lambda$  and  $\Lambda$  and calculate the values of the  $Q$ -statistics.
4. Obtain the  $\lambda$  and  $\Lambda$  that minimize the  $Q$ -statistic.

Note that instead of picking any range of values of  $\lambda$  and  $\Lambda$ , we select the range from the initial estimation. This can reduce the computation considerably. Restricting

the search to estimates relatively close to the true values is important since if  $\lambda$  and  $\Lambda$  are selected too far away from the true values, both  $\hat{\rho}_{1k}$  and  $\hat{\rho}_{2k}$  will tend to 0 very quickly as  $k$  increases. This will incorrectly result in the  $Q$  value being close to 0.

Table 2 shows the results for the data(Figure 5(c)) from model (6), i.e.,

$$X(t) = A_1 \cos(2\pi\beta_1 \frac{(t + \Lambda)^\lambda - 1}{\lambda}) + A_2 \cos(2\pi\beta_2 \frac{(t + \Lambda)^\lambda - 1}{\lambda}) + a(t),$$

where  $A_1 = 10$ ,  $A_2 = 5$ ,  $\beta_1 = 5$ ,  $\beta_2 = 4.5$ ,  $\lambda = 0.5$ ,  $\Lambda = 100$  and  $a(t) \sim N(0,1)$ , and data(Figure 24(c)) from model (29), i.e.,

$$(t^{1/2}D - \alpha_1)(t^{1/2}D - \alpha_2)X(t) = (t^{1/2}D - \beta)a(t),$$

where  $\alpha_1 = -0.0101 - 0.506i$ ,  $\alpha_2 = -0.0101 + 0.506i$ ,  $\beta = 2.375$ ,  $\Lambda = 100$  and  $\text{var}(a(t)) = 0.188$ . When the cyclical behavior is very clear, as in Figure 5(c), this method gives us the very accurate estimates  $\hat{\lambda} = 0.5$  and  $\hat{\Lambda} = 100$ . When elongation of the cyclical behavior of the data is a little weak such as in Figure 24(a), this method still provides good estimates,  $\hat{\lambda} = 0.6$  and  $\hat{\Lambda} = 52$ . Since the  $Q$ -value for  $\lambda = 0.4$  and  $\lambda = 0.5$  are very close, in practice, we would typically pick  $\hat{\lambda} = 0.5$ , instead of  $\hat{\lambda} = 0.4$  for ease of interpolation, i.e. we use the square root time transformation, The corresponding estimator of  $\Lambda$  is then  $\hat{\Lambda} = 100$ , which is exactly equal to the true value.

The approach we recommend here provides a neighborhood of the true value rather than one specific value of  $\lambda$  or  $\Lambda$ . It actually gives us freedom to interpret the final model we obtain. For example, if  $\hat{\lambda} = 0.01$ , we typically use  $\hat{\lambda} = 0$ , which is the logarithmic time transformation, and is easily interpreted.

Table 2: Minimum  $Q_s$  for each Possible  $\lambda$

$\lambda$	Data from Model (6)		Data from Model (29)	
	$\Lambda$	$Q$	$\Lambda$	$Q$
0.0	360	0.017937059	372	0.7521601
0.1	310	0.017526286	320	0.7473886
0.2	250	0.036805857	264	0.7396110
0.3	200	0.015139636	208	0.7162618
0.4	150	0.001666056	154	0.6579882
0.5	100	0.001132114	100	0.5991000
0.6	50	0.014224997	52	0.5314034
0.7	10	0.133403316	9	1.1090868

## 9 Application to Actual Data

We will apply the methods mentioned above to two data sets in this section. One is a geophysical data set called MNTA, and the other is a bat signal called ABEND.

### 9.1 Data MNTA

The original MNTA data set provided by Dr. Brian Stump has 4096 measurements. Figure 26(a) shows the entire data set. Note that the data values at the beginning and at the end of the data set are very close to zero. Therefore, we will only use data values 901-1600 (Figure 26(b)) for analysis. We shall refer to this subset of MNTA as MNTASUB. Figure 26(c) and Figure 26(d) show the usual AR spectrum and sample periodogram of the data, respectively. They both show that the data

has dominant frequencies located in a neighborhood of  $f = 0.03$ . Obviously, the data are compacting over time, and the frequency of the data at the beginning is lower than at the end. Based on the peak detection method discussed in Section 3, we obtained the empirical instantaneous period of the data (Figure 27(a)), where it can be seen that the period is monotonically decreasing over time. From Figure 27(b), the ACF,  $\hat{\rho}_1$ , (“1”) of the first 450 data points vs. the ACF,  $\hat{\rho}_2$ , (“2”) of the remaining data points of MNTASUB, indicates that the two ACFs have different frequencies. Thus, stationarity of the data is questionable. Therefore, we need to test  $H_0 : \lambda = 1$ . Figure 27(c) provides a diagnostic display for testing  $\lambda = 1$ , where the 95% acceptance region of  $\hat{\rho}_2$  under the null hypothesis is in the region between the two solid lines. It seems that the region does not contain the observed  $\hat{\rho}_2$  very well, especially for the first 40 lags. Figure 27(d) shows the empirical distribution of the  $Q$ -value under the null hypothesis, the 95th-quantile ( $Q_{.95}$ ) with the dashed line of the  $Q$ -value, and the observed  $Q$  (arrow). The observation is obviously larger than  $Q_{0.95}$ , and we therefore reject the null  $H_0 : \lambda = 1$ . Thus traditional Fourier spectral analysis is not appropriate. Currently, many methods such as wavelet analysis, short window Fourier analysis (Percival and Walden 2000) and autoregressive models with time-varying coefficients (Rajan and Rayner, 1996) are used to analyze data with time-varying frequencies. Figure 28 shows the result of short window Fourier analysis, where the x-axis is frequency, y-axis is time, and the dark line represents the dominant frequency at each time point.

From the plot, the data appear to have monotonically increasing dominant frequency. We will therefore consider a  $G(\lambda)$ -stationary model with  $\lambda > 1$  for the data. To identify  $\lambda$  and  $\Lambda$ , we set initial ranges of  $\lambda$  and  $\Lambda$  from 1 to 3 and from 0 to 400, respectively. The initial estimation gives the range of  $\lambda$  from 2.0 to 2.8. From Table

3, the value of  $\lambda$  appears to be located in a neighborhood from 2.4 to 2.8. We then pick  $\hat{\lambda} = 2.5$  and  $\hat{\Lambda}$  is 216. The corresponding discrete dual of MNTASUB(Figure 30(a)) seems to have a constant frequency over the entire time period. Figure 30(b) and Figure 30(c) show the AR spectrum and sample periodogram. The discrete dual data has a low dominant frequency located around  $f = 0.022$ , and also it has sharper peak than that of the original data set shown in Figure 26. The empirical instantaneous period (Figure 30(d)) of the discrete dual is reasonably flat. Figure 29 shows the sample instantaneous spectrum based on  $\lambda = 2.5$  and  $\Lambda = 216$ . Comparison with the time-frequency plot in Figure 28 shows that when  $t < 200$  and  $t > 750$ , the sample instantaneous spectrum clearly identifies the dominant frequency of the data, while the short window Fourier analysis does not. For  $200 < t < 750$ , both methods give similar results. Next, we fit a  $G(p, q; \lambda)$  model to the data and obtain a  $G(15, 0; 2.5)$  model with offset 216. When sampling the data at  $G(\lambda)$ -time scale with the  $G(\lambda)$ -sampling interval  $\Delta = 18203.7$ , the discrete dual is given by

$$\begin{aligned}
& (1 - 1.97B + 0.99B^2)^2(1 + 0.7611B + 0.7626B^2)(1 - 0.6179B + 0.7472B^2) \\
& (1 + 0.0734B + 0.7247B^2)(1 + 1.2655B + 0.7099B^2) \\
& (1 + 1.5611B + 0.6651B^2)(1 - 0.7476B)Y_u = \varepsilon_u,
\end{aligned} \tag{30}$$

where  $\text{var}(\varepsilon_u)=3.361$ . An  $\text{AR}(p)$  model will be used as a competing model to evaluate the forecast performance of  $G(15;2.5)$  model. With the highest order set at 30, the AIC criterion selects an  $\text{AR}(22)$  model for MNTASUB. Figure 31 compares the forecast performances of the  $\text{AR}(22)$  model and the  $G(15;2.5)$  model. Clearly, the  $G(15;2.5)$  model has much better forecasts than  $\text{AR}(22)$  as would be expected. Table 4 shows the MSE of forecasts at different lags using the  $\text{AR}(22)$  and the  $G(15;2.5)$  model, respectively. The latter model almost reduces the MSE by 80% forecasts for

Table 3: Minimum Qs for each Possible  $\lambda$  for Data MNTASUB

$\lambda$	offset	Q	slope	t
2.0	8	4.708274	0.0100409943	1.9020885
2.1	48	3.599942	0.0104302499	2.1193730
2.2	88	2.845673	0.0109098621	2.3429982
2.3	128	2.325702	0.0114238232	2.5667171
2.4	176	1.968625	0.0111080888	2.6395371
2.5	216	1.675376	0.0116661060	2.8579594
2.6	256	1.479355	0.0122018868	3.0689614
2.7	304	1.310339	0.0120100714	3.1343856
2.8	360	1.250445	0.0112107764	3.0593131

at lag 10 to lag 60. Figure 32 shows the instantaneous spectrum for MNTASUB based on the  $G(15; 2.5)$  model, which is almost the same as the sample instantaneous spectrum shown in Figure 29. From Figure 32, the dominant frequency of the original data is about  $f = 0.014$  at  $t = 200$  and about  $f = 0.041$  at  $t = 845$ , which implies that the data has an instantaneous period of 70 at the 200th data point and an instantaneous period of 24 at the 850th data point. From Figure 33, we can see that these observations are consistent with the original data set.

## 9.2 Data ABEND

The bat signal is recorded from the *Nyctalus noctula* bat. One can download the data from the web site: [http://www.zwergfledermaus.de/wav/wave\\_e.htm](http://www.zwergfledermaus.de/wav/wave_e.htm). The data set has 82,958 measurements, and the sample rate is about 25kHz. In this section,

Table 4: MSE of Forecasts: AR(22) vs G(15,0;2.5) for Data MNTASUB

lags	AR(22) Model	G(15, 0; 2.5) Model	Improvement
10	3164788	493374	84%
20	7524439	831502	89%
30	5767593	665691	88%
40	4651343	570987	88%
50	3905185	808824	79%
60	3404564	855499	75%

we only analyze a typical subset, referred to as ABENDSUB, of length 280. From Figure 34(a), ABENDSUB has cyclical behavior elongating in time. Figure 34(b) shows its uninformative sample spectrum, which has many peaks, namely, 0.088, 0.1139, 0.2649, 0.4079. Separating ABENDSUB into two parts of length 140 each and calculating the ACF for each part, the two ACFs are quite different (Figure 34(d)). Therefore, the stationarity of the data is questionable. We test the null hypothesis that the data are from a stationary process using the bootstrap method, and in Figure 35(a), we see that the observation  $Q$  is larger than  $Q_{.95}$ , the 95th quantile of the empirical distribution of  $Q$ -values under the null, so we reject the null.

From the data in Figure 34(a) and the empirical instantaneous period in Figure 34(c), it is clear that the periods are increasing in time. Fitting a  $G(\lambda)$ -stationary model with  $\lambda < 1$  for the data, we set the initial range of  $\lambda$  and  $\Lambda$  from 0 to 1 and from 0 to 250, respectively. From Table 5, we see that  $\hat{\lambda} = 0$  and  $\hat{\Lambda} = 188$  are selected as the best parameters estimates. That is, we fit an M-stationary process to the data.

The corresponding dual data, the sample spectrum of the dual and the empirical instantaneous period are shown in Figures 35(b)-(d), respectively. The dual data has strong cyclical behavior, with constant period. The sample spectrum shows three frequencies, i.e., 0.12, 0.24 and 0.36. Figure 36 and Figure 37 show the instantaneous spectrum under the assumption the data are from an M-stationary process and the time-frequency plot generated using a window based Fourier method. These plots provide similar information concerning the time varying frequency behavior. However, the sample instantaneous spectrum clearly shows three dominant frequencies in the data, while the time-frequency plot only gives the most dominant one.

By fitting an AR( $p, q$ ) model to the discrete dual, we obtain an Euler(12) model for the data. When the sample rate of the Euler time scale is  $h = 1.00326$ , its discrete dual is given by

$$\begin{aligned}
& (1 - 1.4448B + 0.9954B^2)(1 - 0.0703B + 0.9732B^2) \\
& (1 + 1.2691B + 0.9193B^2) \\
& (1 - 1.7282B + 0.7599B^2)(1 - 0.7885B + 0.6537B^2) \\
& (1 + 1.502B + 0.5758B^2)(Y_u + 0.01636) = \varepsilon_u, \tag{31}
\end{aligned}$$

where  $\text{var}(\varepsilon_u) = 0.002722$ . Figure 38 shows the instantaneous spectrum based on the Euler(12) model. Since the data have multiple frequencies, it is not easy to evaluate the model-based instantaneous spectrum by visual methods as we did in Figure 33. However, when we compare it with the previous sample instantaneous spectrum and the time-frequency plot, the instantaneous spectrum based on the Euler(12) model characterized the change of the three dominant frequencies of ABENDSUB very well. Based on model (31), Figures 39(a)-(d) show the snapshots of the Euler(12) spectrum at four specific times, i.e.,  $t = 1$ ,  $t = 80$ ,  $t = 160$  and  $t = 220$ . Note that since the

Table 5: Minimum Qs for each  $\lambda$  for Data ABENDSUB

$\lambda$	offset	Q	slope	t
0.0	188	0.07823766	0.001953457	1.631810
0.1	160	0.09287909	0.001846839	1.514361
0.2	128	0.14823380	0.002247002	1.783700
0.3	96	0.19932705	0.002878826	2.166317
0.4	64	0.35008821	0.003992134	2.727323
0.5	36	0.65244955	0.005119565	2.990009
0.6	12	1.50383413	0.006702452	2.960328
0.7	0	8.93196774	0.004249245	1.500202

frequency monotonically decreases with time, the highest dominant frequency is so fast that it cannot be identified at the sample rate used for  $t < 25$ . To evaluate the forecast performance of the Euler(12), we use an AR( $p$ ) as a competing model. With the highest order set at 20, the AIC criterion picks an AR(16). Table 6 shows the MSE of forecasts at different lags, where it can be seen that Euler(12) reduces the MSE by 90% over the forecasts for the AR(16) model for all lags from lag 10 to lag 60. Therefore, from both spectral analysis and forecast performance points of view, the Euler(12) is a good model for ABENDSUB.

## 10 Conclusion

The goal of this article is to extend the work of Gray and Zhang(1988), Gray, Vijverberg and Woodward(2004) and Choi (2003) concerning the continuous M-stationary

Table 6: Forecasts Performance for Data ABENDSUB: AR(16) vs Euler(12).

lags	AR(16) Model	Euler(12) Model	Improvement
10	0.086352	0.013884	84%
20	0.221560	0.016544	93%
30	0.387166	0.020579	95%
40	0.341677	0.018365	95%
50	0.452858	0.040770	91%
60	0.426750	0.033221	92%

process and the discrete M-stationary process. A continuous M-stationary process can be transformed to a continuous weakly stationary process, referred to as the dual process, through a logarithmic time transformation. The M-stationary process provides a better fit for the non-stationary data with cycle length increasing approximately linearly in time than the usual methods based on a stationarity assumption. The  $G(\lambda)$ -stationary processes are proposed to model data with a wide range of time-frequency behavior, especially for the processes with frequencies that monotonically increase or decrease over time. The usual stationary processes and M-stationary processes are two special cases of the  $G(\lambda)$ -stationary process, i.e.,  $\lambda = 1$  and  $\lambda = 0$ , respectively. Using the Box-Cox transformation on the time scale, the  $G(\lambda)$ -stationary processes is transformed to a stationary dual. Most properties of stationary processes and M-stationary processes are special cases of the corresponding properties of  $G(\lambda)$ -stationary processes. The general instantaneous period and general instantaneous frequency are introduced to measure the change of the period and the frequency in  $G(\lambda)$ -stationary processes. These concepts are first introduced

by Gray, Vijverberg and Woodward (2004) for measuring the change of the period of the data for M-stationary processes. Here we give a more general definition. The general instantaneous period measures the length of the next cycle starting at any time. The general instantaneous frequency is its reciprocal.

A typical class of  $G(\lambda)$ -stationary process, the  $G(p, q; \lambda)$  processes, are proposed. When  $q = 0$ , this is referred to as the  $G(p; \lambda)$  process. Their dual processes are the continuous ARMA( $p, q$ ) processes and the continuous ARMA( $p$ ) processes, respectively. The continuous processes can only be applied to real data through discretization. Extending the theorem by Phadke and Wu (1974), we prove that sampling a  $G(p; \lambda)$  or  $G(p, q; \lambda)$  process at the  $G(\lambda)$ -time scale results in a discrete ARMA( $p, r$ ) process, where  $r \leq p - 1$  if the Nyquist frequency is greater than the highest frequency of the corresponding  $G(p; \lambda)$  or  $G(p, q; \lambda)$  process. The importance of the theorem is that inference based on the sampled data can be made regarding the underlying continuous process. Moreover, in Phake and Wu's research, the coefficients of the discrete stationary process are restricted to a specific region. In this thesis, we extend their method and find a unique continuous stationary or  $G(\lambda)$ -stationary for each discrete stationary process.

For the  $G(\lambda)$ -stationary processes, we use the instantaneous spectrum to provide a spectral representation that describes the manner in which the frequency changes over time. Time-frequency analysis has been approached in the literature using the window-based Fourier transform and wavelets. The instantaneous spectrum analysis has clearer and smoother results, especially, for data with multiple frequencies changing over time when the change can be described by a monotonic transformation.

In practice, most data are collected at equally spaced time points. Methods to simulate equally spaced realizations from a continuous  $G(p, q; \lambda)$  model, model the

equally spaced data, and forecast are introduced in this article. For the latter two procedures, the main idea is to interpolate the data at the  $G(\lambda)$ -time scale from the equally spaced data and reinterpolate forecasts based on interpolated data to obtain the forecasts on the equally spaced time scales.

A challenging problem in dealing with the  $G(\lambda)$ -stationary process in this research is to estimate the value of  $\lambda$  and offset  $\Lambda$ , which affect both the instantaneous spectrum and the final model. A bootstrap method is used to test the null hypothesis that the data are from a stationary process. Also, we provide a procedure to narrow the range of possible values of  $\lambda$  and  $\Lambda$  and to identify the best model based on the ACF of the dual data. The sample spectral density may be used instead of the ACF in the estimation procedure, and this is a topic for future research.

Finally, the  $G(p, q; \lambda)$  models are used to fit the two actual data, i.e., Data MNTA-SUB and Data ABNEDSUB, respectively. From the spectral analysis point of view, the  $G(p, q; \lambda)$  models do a good job of describing the frequencies and the change of frequency over a given time interval for each data set. Also the  $G(p, q; \lambda)$  models outperform the usual AR models in forecast performance.

This research marks the beginning of a new area in non-stationary analysis. A natural extension of this research is to explore processes with several different  $\lambda$ 's. This a very fruitful area of research with much work left to be done.

## References

- [1] Bartlett, M. S. (1946), "On the Theoretical Specification and Sampling Properties of Autocorrelated Time Series", *Journal of the Royal Statistical Society*, B8, 27.
- [2] Boashash, B.(1992), "Estimating and Interpreting the Instantaneous Frequency

- of a Signal-Part; Part I: Fundamentals, II: Algorithms”, *Proceedings of the IEEE*, 80, 4, 519-569.
- [3] Choi, E., Gray, H. L. and Woodward, W. A. (2003), “Euler( $p, q$ ) Processes with Spectral Analysis of Non-stationary Time Series”, *SMU Tech Report*, #313.
- [4] Gradsheyn, I. S and Ryzhik, I. M. (2000), *Mellin Transform*, CA: Academic Press.
- [5] Gray, H. L., and Zhang, N. (1988), “On a Class of Nonstationary Processes”, *Journal of Time Series Analysis*, 9,2, 133-154.
- [6] Gray, H. L. and Vijverburg, C. P.(2002), “Time Deformation and The M-spectrum—A New Tool for Cyclical Analysis”, *SMU TECH Report*, #314.
- [7] Gray, H. L., Vijverburg, C. P., and Woodward, W. A.(2004), “Nonstationary Data Analysis by Time Deformation”, *Communication in Statistics A*, in press.
- [8] Hannan, E. J. (1965), *Group Representation and Applied Probability*, Methuen’s Review Series in Applied Probability, Vol. 3. London: Methuen.
- [9] Hosking, J. R. M. (1981), “Fractional Differencing”, *Biometrika*, 68, 165-76.
- [10] Jenkins, G. M. and Watts, D. G. (1968), *Spectral analysis and its applications*, San Francisco: Holden-Day.
- [11] Percival, D. B. and Walden, A. T. (2000), *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge UK.
- [12] Phadke, M. S. and Wu, S. M. (1974), “Modeling of Continuous Stochastic Processes from Discrete Observations with Application to Sunspots Data”, *Journal of the American Statistical Association*, 69, 325-329.

- [13] Priestley, M. B. (1981), *Spectral Analysis and Time Series*, Vols. 1 and 2, Academic Press, New York.
- [14] Rajan, J.J. and Rayner, P.J.W. (1996), “Generalized feature extraction for time-varying autoregressive models”, *Signal Processing , IEEE Transactions on*, 44, 10, 2498-2507.
- [15] Ripley, B. D. (1977), “Modelling Spatial Patterns”, *Journal of the Royal Statistical Society. Series B (Methodological)*, 39, 2, 172-212.
- [16] Stock, J. (1987), “Measuring Business Cycle Time”, *Journal of Political Economy*, 95, 6, 1240-1261.
- [17] Stock, J. (1988), “Estimating Continuous — Time Processes Subject to Time Deformation: An Application to Postwar U.S. GNP”, *Journal of the American Statistical Association*, 83, 401, 77-85.
- [18] Tsay, R. S. (1992), “Model Checking via Parametric Bootstraps in Time Series Analysis”, *Applied Statistics*, 41, 1, 1-15.

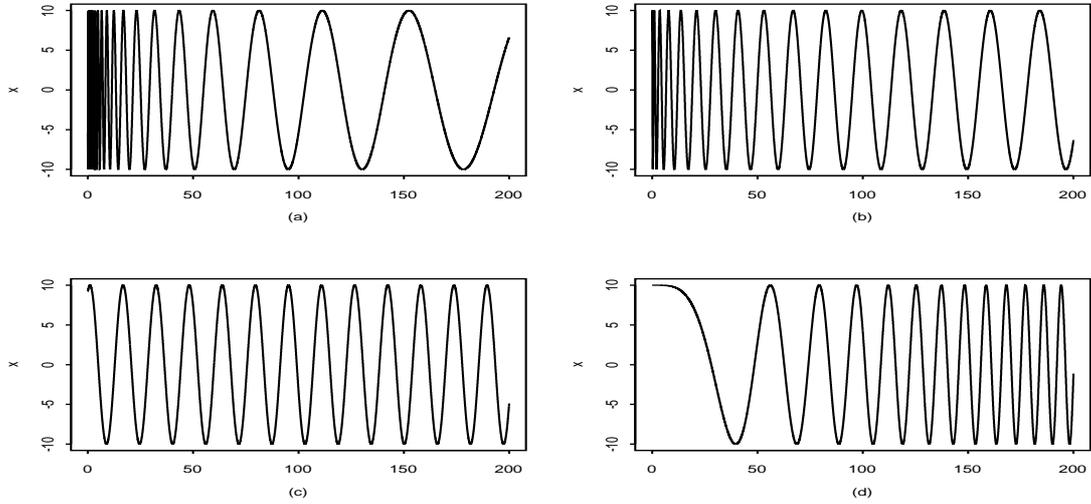


Figure 1: Realizations from the model in Example 2.1: (a)  $A=10$ ,  $\beta = 20$ ,  $\phi = 0$ ,  $\lambda = 0$ ; (b)  $A=10$ ,  $\beta = 3.5$ ,  $\phi = 0$ ,  $\lambda = 0.5$ ; (c)  $A=10$ ,  $\beta = 0.4$ ,  $\phi = 0$ ,  $\lambda = 1$ ; (d)  $A=10$ ,  $\beta = 0.004$ ,  $\phi = 0$ ,  $\lambda = 2$ .

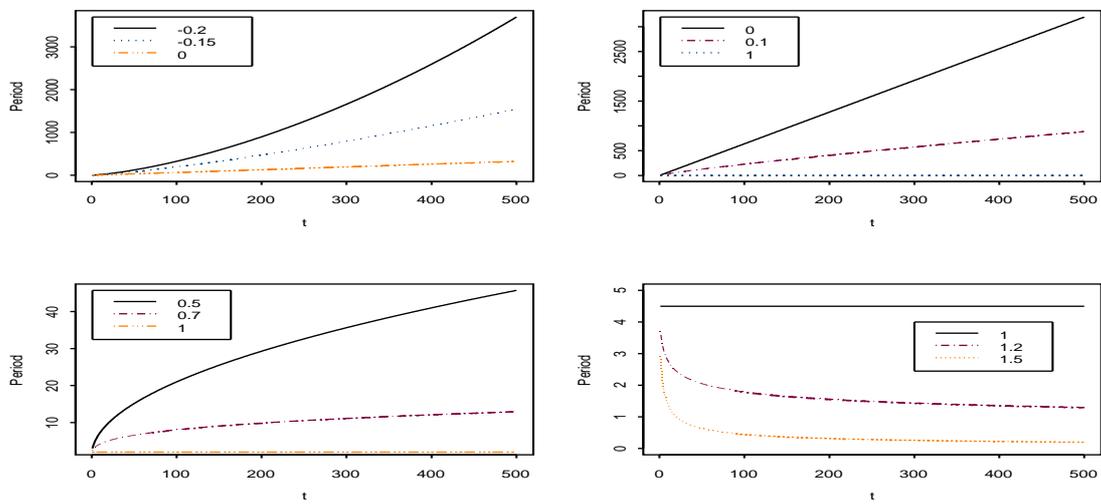


Figure 2: The Plot of the Instantaneous Periods for Different  $\lambda$ 's

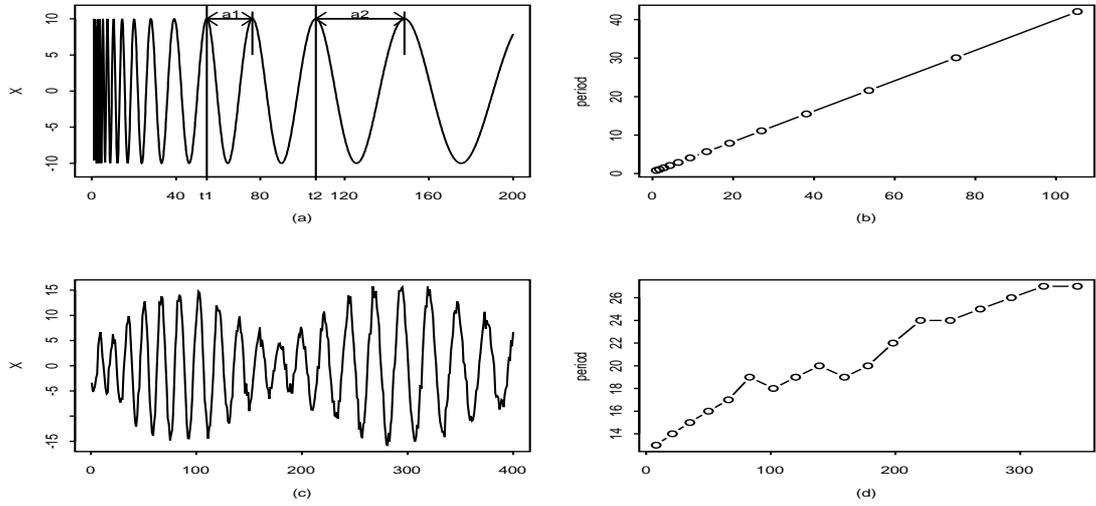


Figure 3: Data and EGIP's from Two Model:(a)The Signal from model (5); (b) EGIP of model (5) vs  $t$ ; (c) A Realization of equation (6); (d) EGIP of model (6) vs  $t$

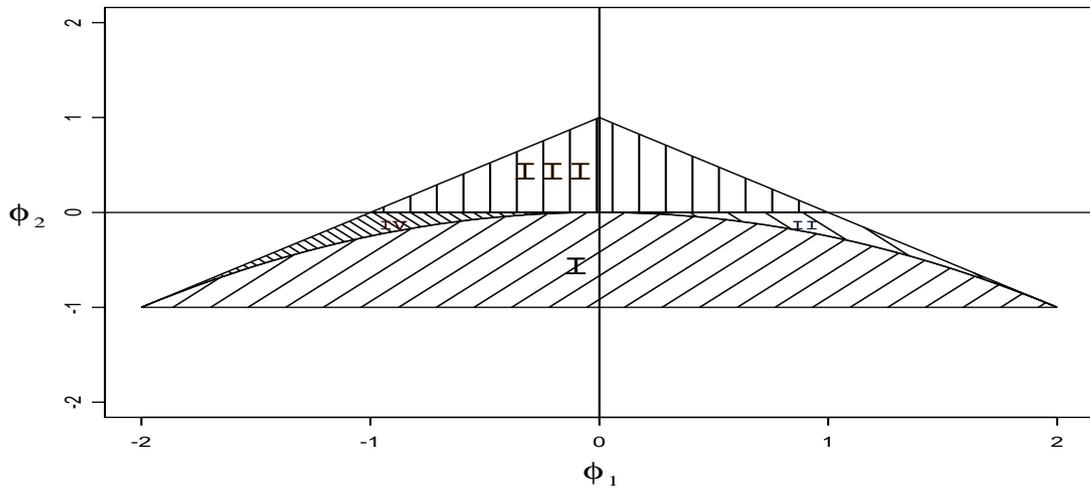


Figure 4: Stationary region covered by  $\phi_1$  and  $\phi_2$  corresponding to discretized continuous ARMA process

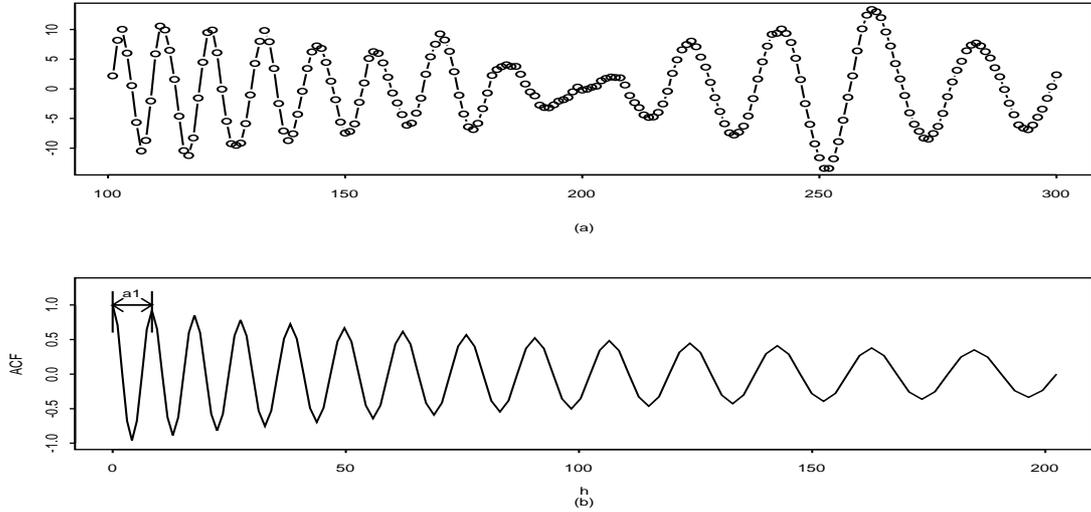


Figure 5: A realization (a) and ACF at time  $t = 101$  (b) of a  $G(2,1;0)$  process with the discrete dual (20) at sampling rate  $h = 1.0055$ .

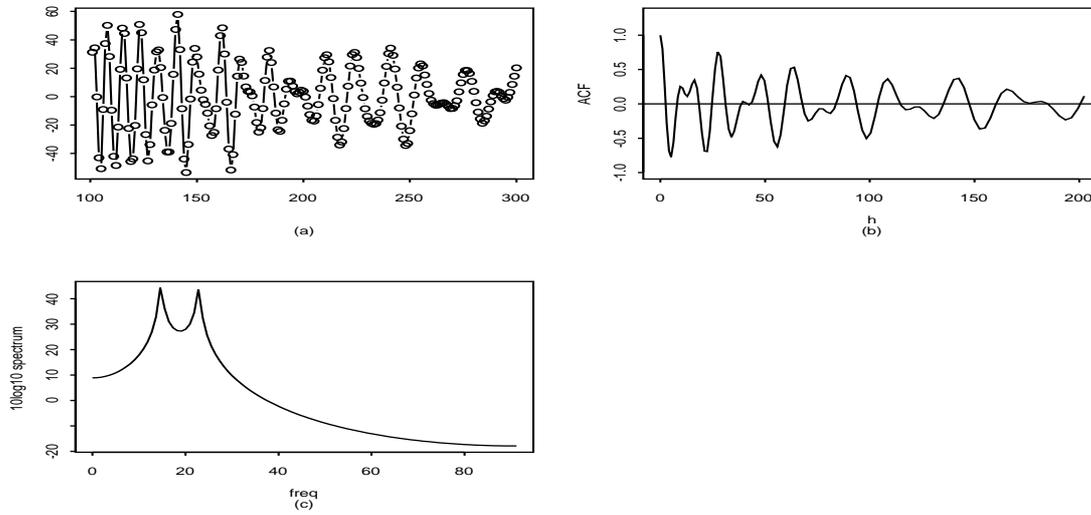


Figure 6: A realization (a), ACF at time  $t = 101$  (b) of a  $G(4,3;0)$  process with the discrete dual in Equation (21) at sampling rate  $h = 1.0055$  and  $G(\lambda)$ -spectrum of the  $G(4,3;0)$  process

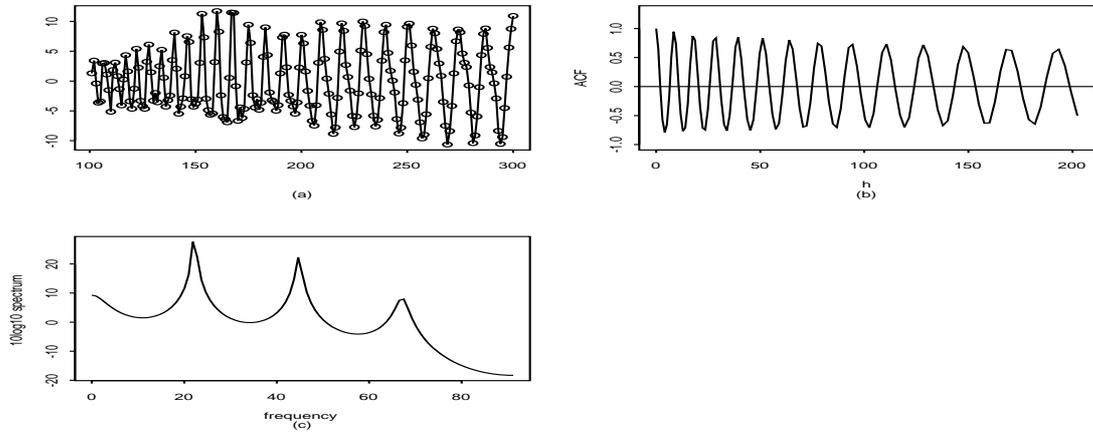


Figure 7: A realization (a), ACF at time  $t = 101$  (b) of a  $G(6,5;0)$  process with the discrete dual in Equation (22) at sampling rate  $h = 1.0055$  and  $G(\lambda)$ -spectrum of the  $G(6,5;0)$  process

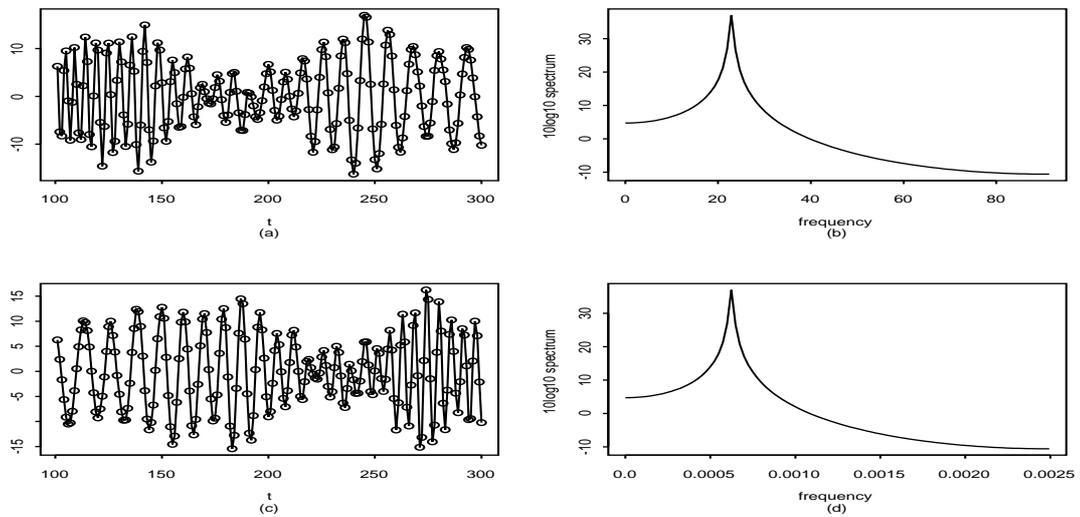


Figure 8: A realization (a),  $G(\lambda)$ -spectrum (b) of a  $G(2,1;0)$  process with the discrete dual (23) at sampling rate  $h = 1.0055$ . A realization (c),  $G(\lambda)$ -spectrum (d) of a  $G(2,1;2)$  process with the discrete dual (23) at sampling rate  $h = 1.0055$ .

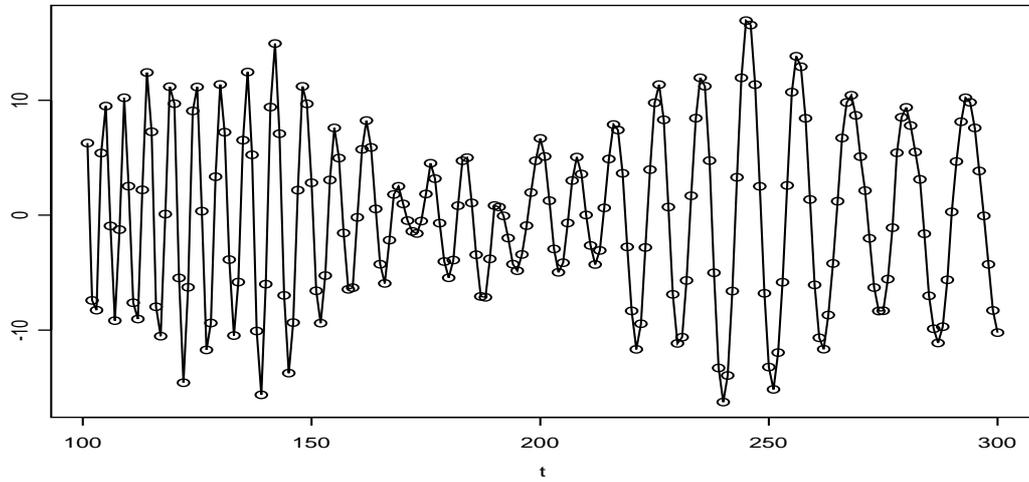


Figure 9: A realization from a  $G(2,1;0)$  model with discrete dual (23) at sample rate  $h = 1.0055$  and  $\text{var}(\varepsilon_k) = 1$ .

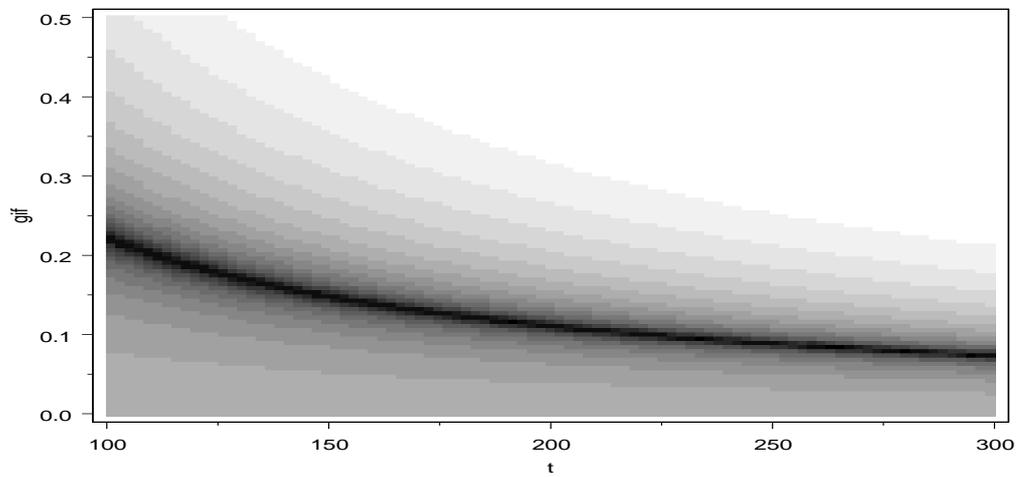


Figure 10: Instantaneous spectrum of the data in Figure 9

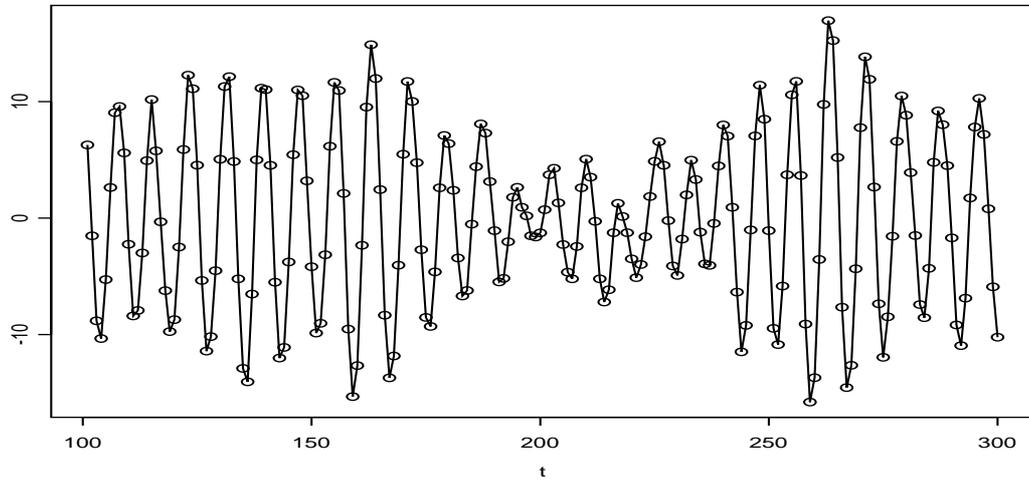


Figure 11: A realization from a  $G(2,1;1)$  model with discrete dual (23) at  $G(\lambda)$  sample interval  $\Delta = 1$  and  $\text{var}(\varepsilon_k) = 1$ .

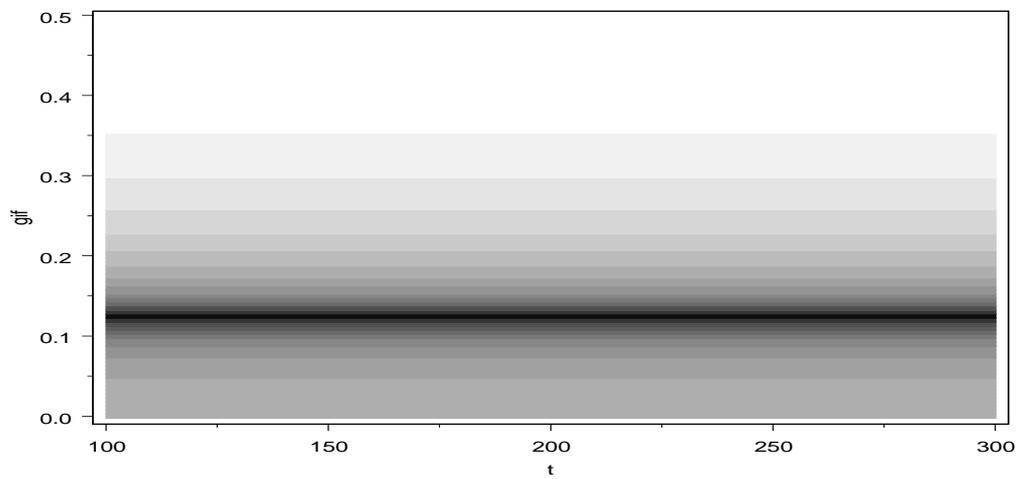


Figure 12: Instantaneous spectrum of the data in Figure 11

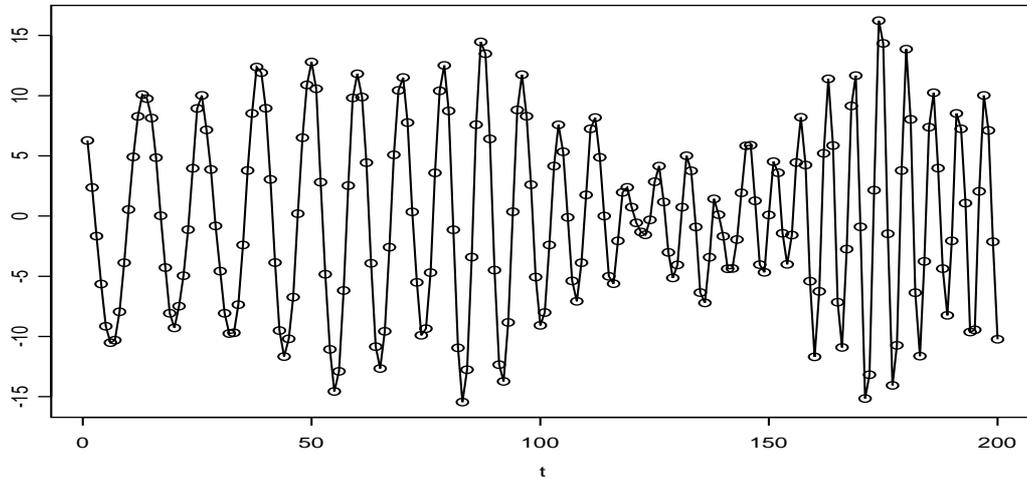


Figure 13: A realization from a  $G(2,1;2)$  model with discrete dual (23) at  $G(\lambda)$  sample interval  $\Delta = 200.5$  and  $\text{var}(\varepsilon_k) = 1$ .

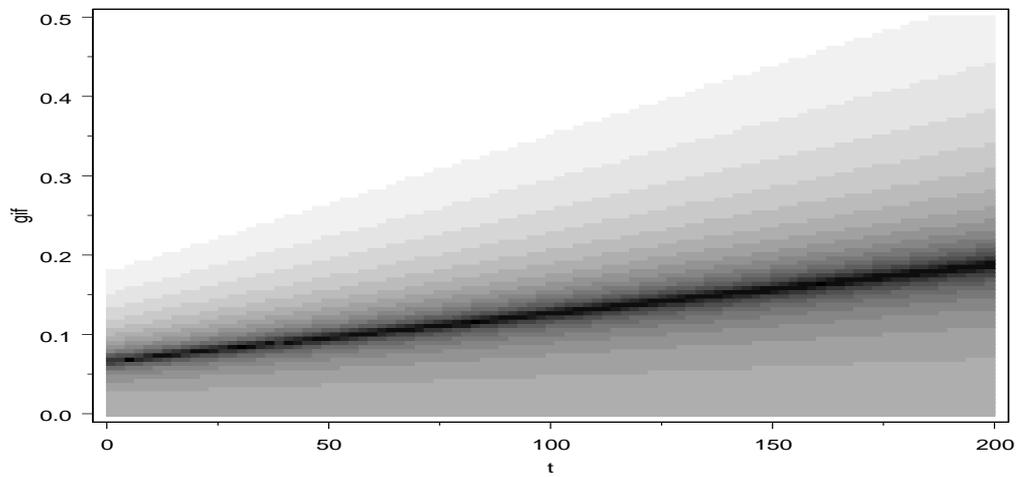


Figure 14: Instantaneous spectrum of the data in Figure 13

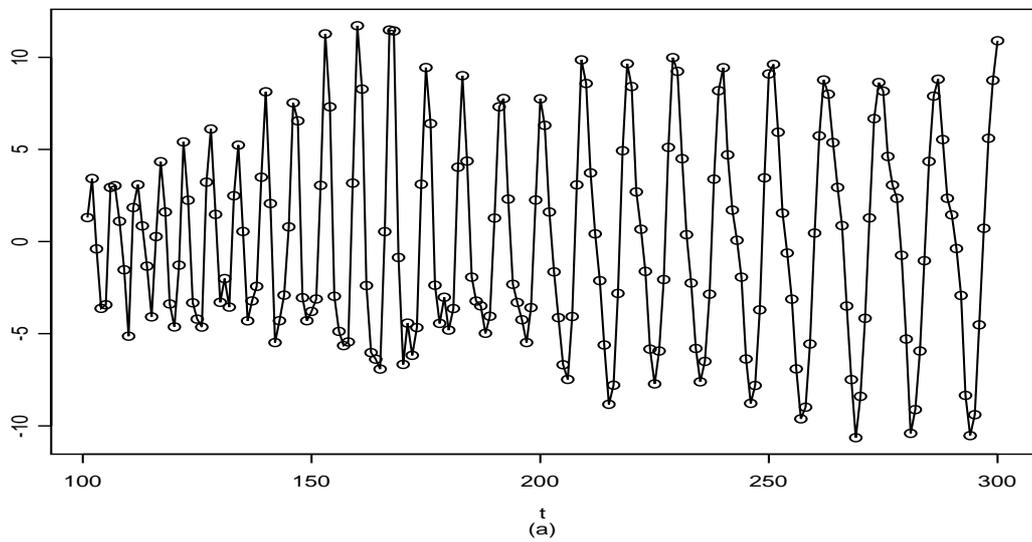


Figure 15: A realization from a  $G(6,5;0)$  model

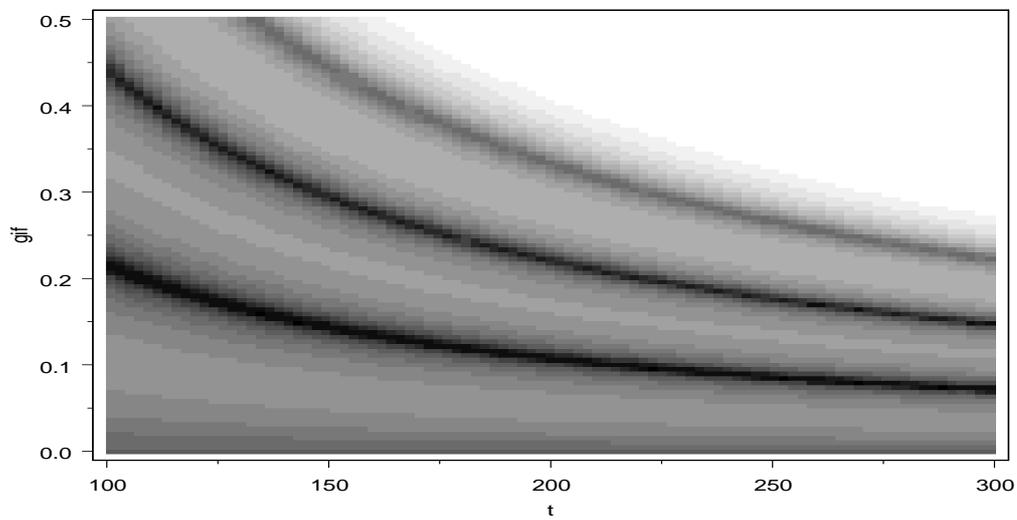


Figure 16: Instantaneous spectrum of the data in Figure 15

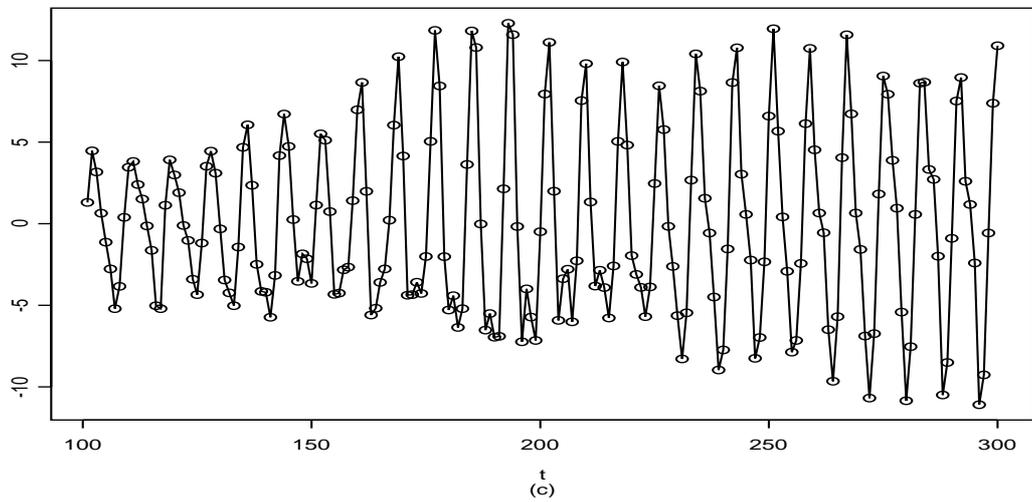


Figure 17: A realization from a  $G(6,5;1)$  model

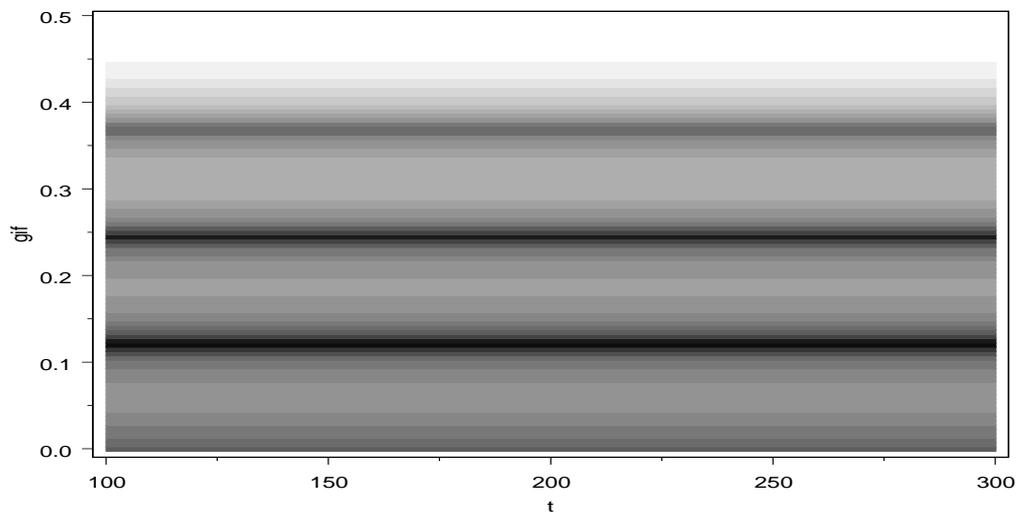


Figure 18: Instantaneous spectrum of the data in Figure 17

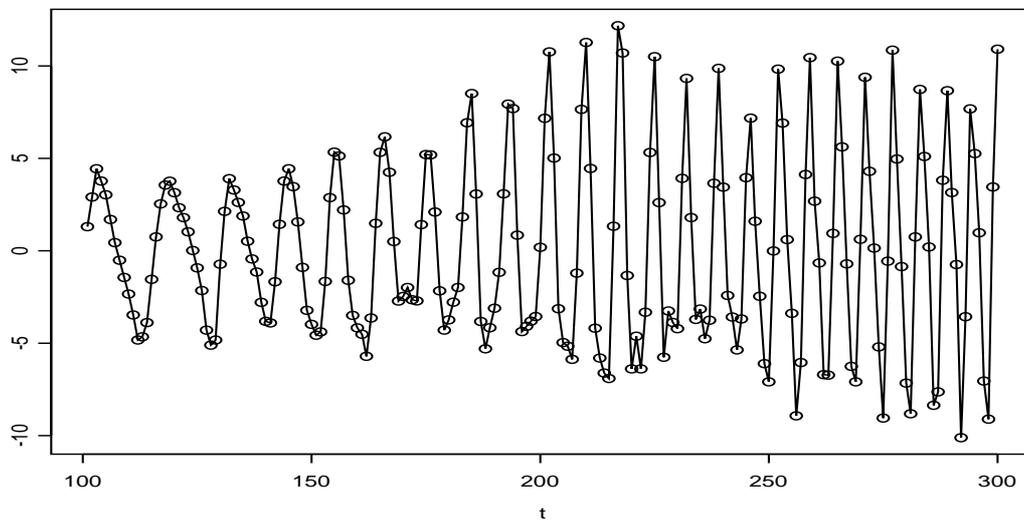


Figure 19: A realization from a  $G(6,5;2)$  model

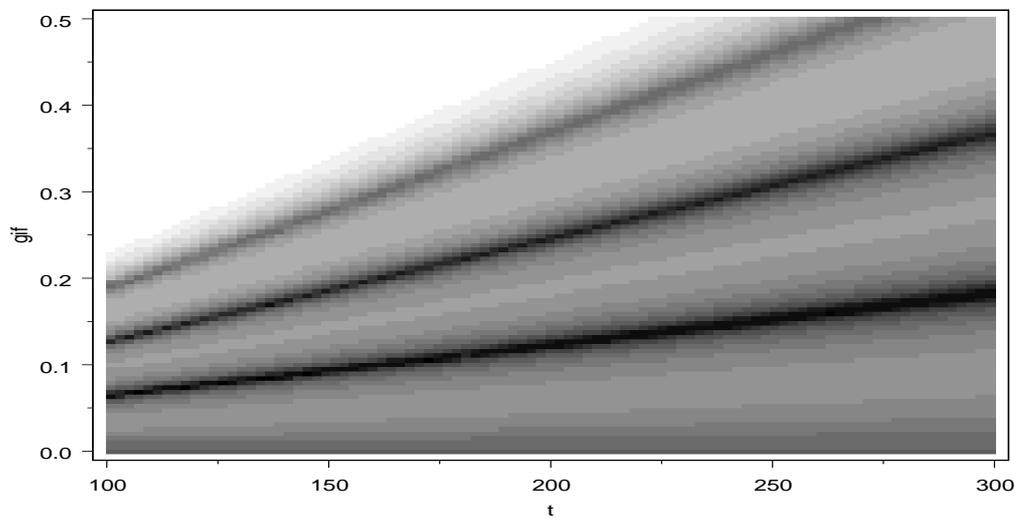


Figure 20: Instantaneous spectrum of the data in Figure 19

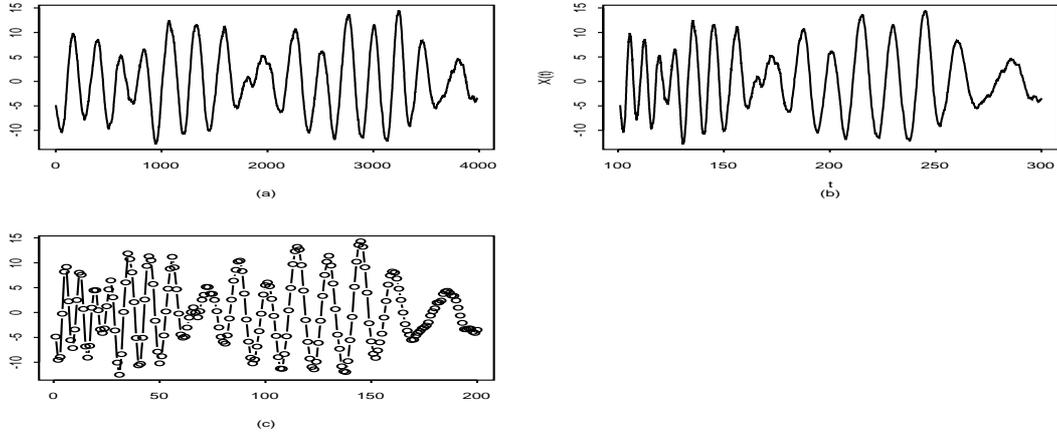


Figure 21: (a) A realization(3981 data points) of model (26); (b) An approximate of a realization of  $X(t)$  for  $101 \leq t \leq 300$ ; (c) An equally spaced data set with length 200 of model (25) given offset  $\Lambda = 100$ .

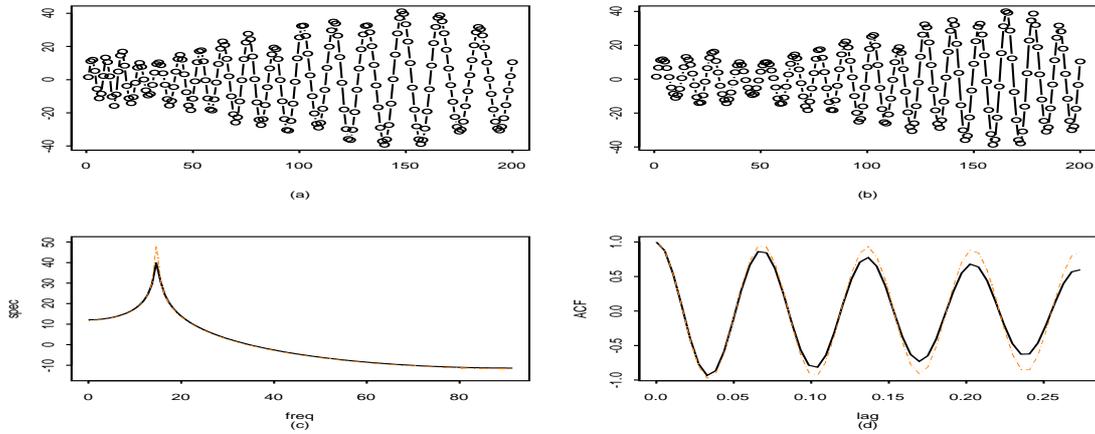


Figure 22: (a) An equally spaced realization with length  $n = 200$  of model (25);(b) The interpolated discrete dual data; (c) The M-spectrum of true model(dashed line) vs that of the based on model (27)(solid line); (d) The M-autocorrelation function of true model(dashed line) vs that of the based on model (25)(solid line).

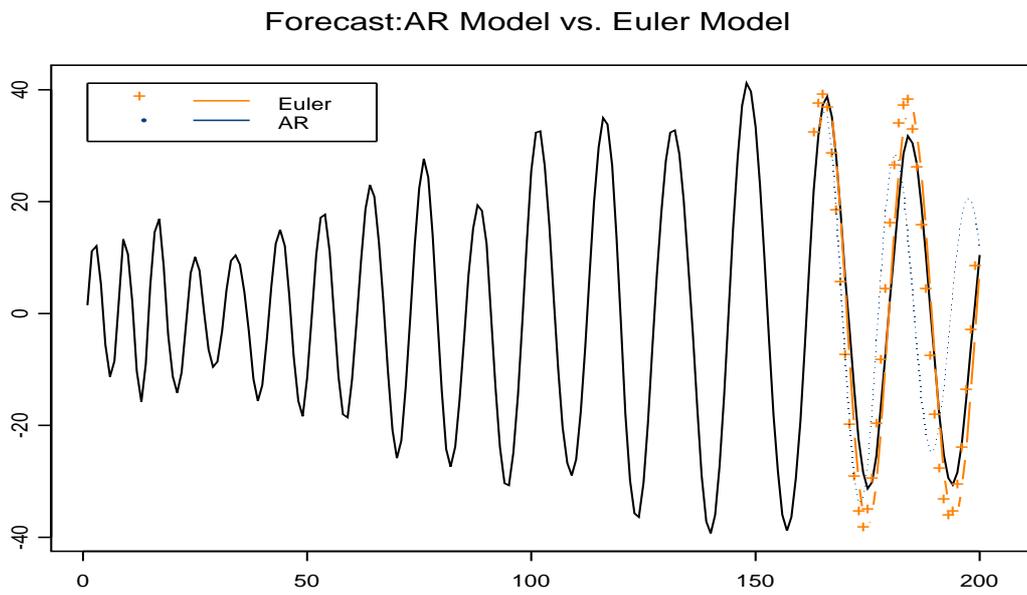


Figure 23:  $l$ -step ahead forecasts based on the Euler model vs  $l$ -step ahead forecasts based on the usual discrete AR model.

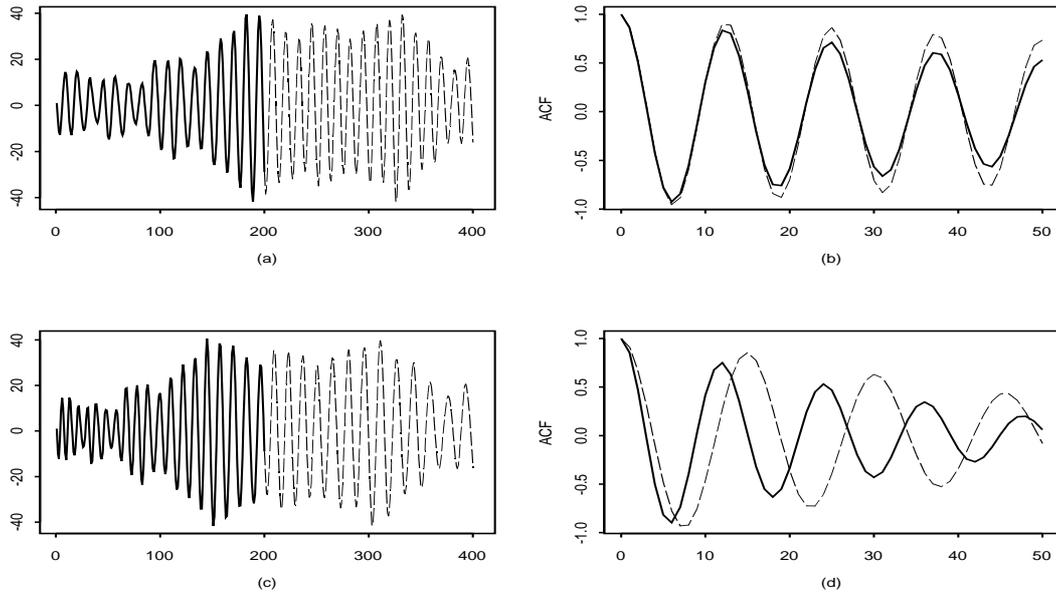


Figure 24: Equally spaced samples of  $G(2,1;\lambda)$  model and their SACF's for  $\lambda = 1$  and  $\lambda = 0.5$ , respectively. The solid lines represent half samples and the corresponding SACF's, while the dot line represent the second half samples and the corresponding SACF's. (a) Equally spaced sample of model (28). (b) SACF of the sample in (a). (c) Equally spaced sample of model (29) with offset=100. (d) SACF of the sample in (c).

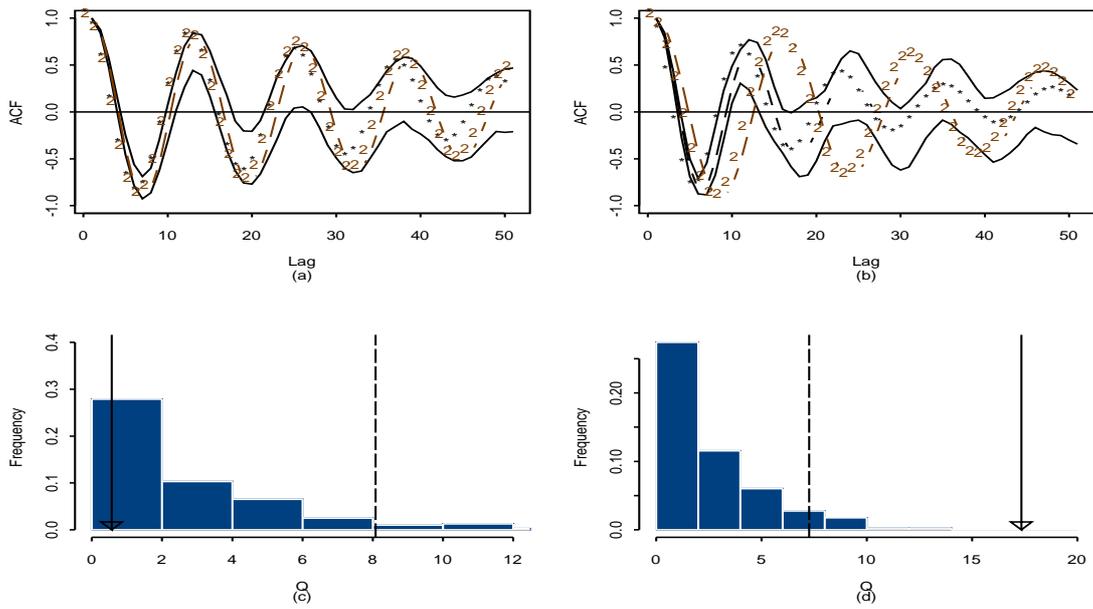


Figure 25: (a)Acceptance envelopes(two solid lines) for SACF of model (28); (b) Acceptance envelope for SACF of the Realization from model (29); (c) Empirical distribution of  $Q$ ,  $Q_{.95}$  (vertically dashed line) and the observation value of  $Q$  for the sample from model (28); (d) Empirical distribution of  $Q$ ,  $Q_{.95}$  (vertically dashed line) and the observation value of  $Q$  for the sample from model (29).

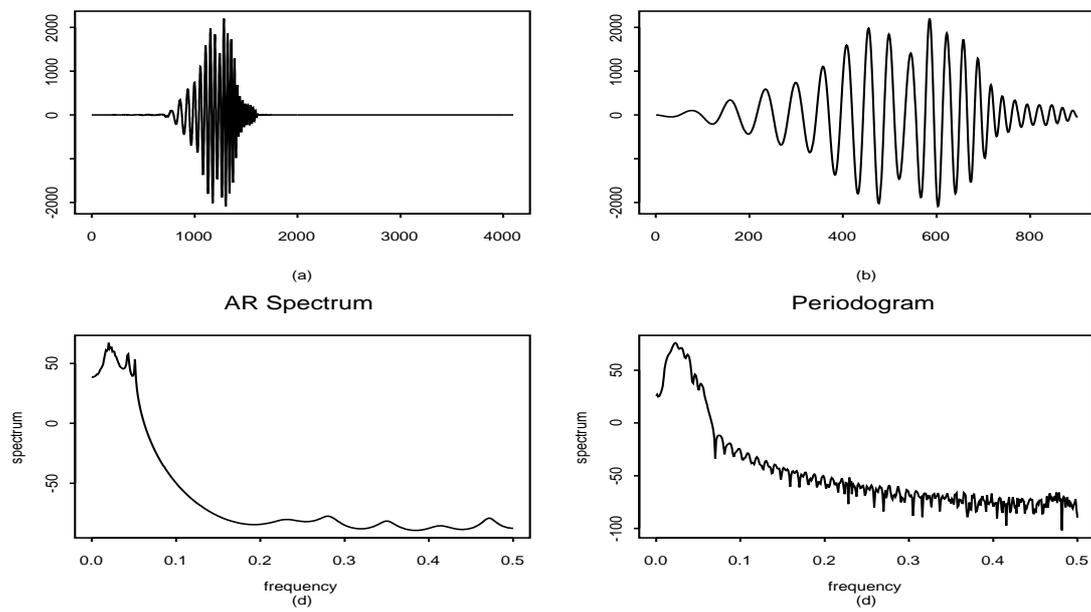


Figure 26: (a) Data MNTA. (b) Data MNTASUB: the subset of MNTA data (from 901th to 1600th) for this analysis. (c) The AR spectrum of Data MNTASUB. (d) The sample periodogram of Data MNTASUB

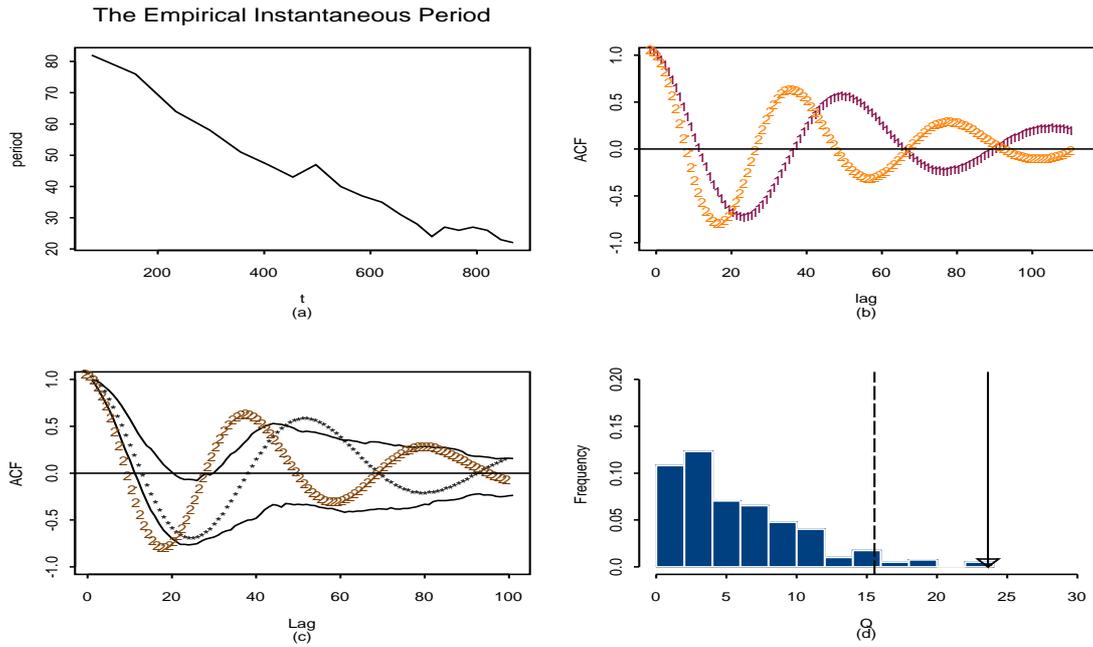


Figure 27: (a) The empirical instantaneous period of Data MNTASUB. (b) the ACF,  $\hat{\rho}_1$ , (the line with sign “1”) of the first 450 data points vs. the ACF,  $\hat{\rho}_2$ , (the line with sign “2”) of the remaining data points of Data MNTASUB. (c) Acceptance envelopes (two solid lines) for  $\hat{\rho}_1$  of DATA MNTASUB. (d) The empirical distribution of  $Q$ ,  $Q_{.95}$  (vertically dashed line) and the observation value (the line with arrow) of  $Q$  for Data MNTASUB.

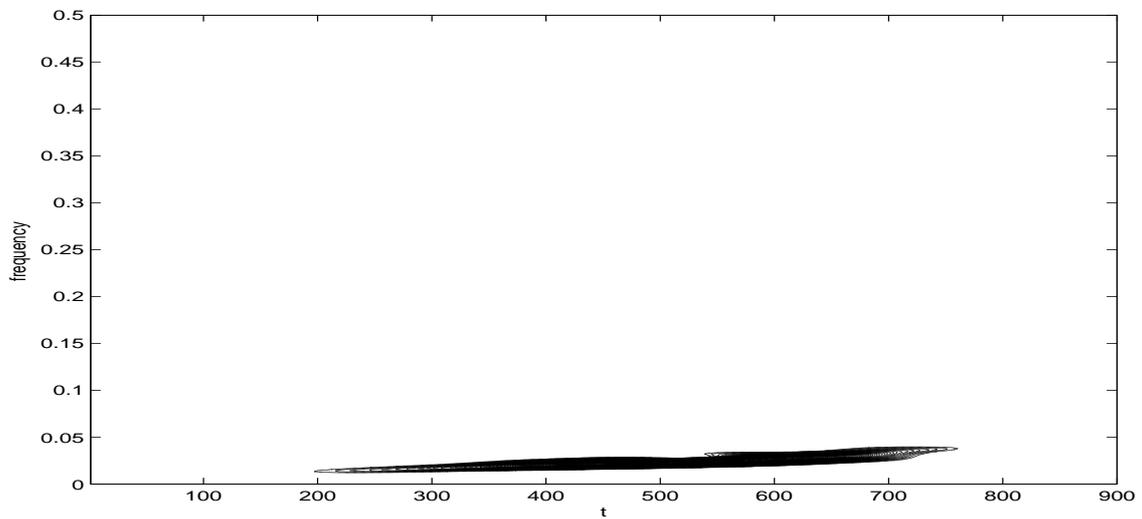


Figure 28: Short-window Fourier analysis for Data MNTASUB

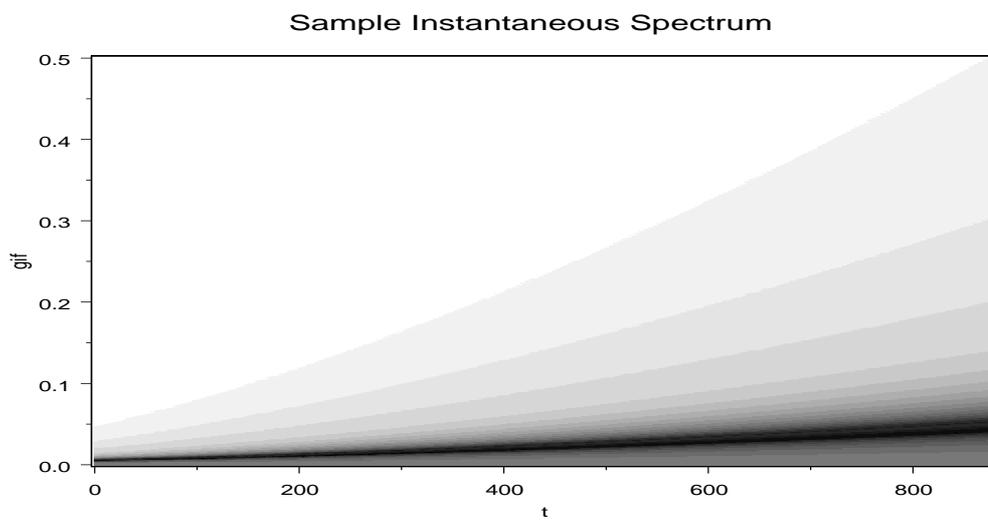


Figure 29: Sample instantaneous spectrum based on  $\lambda = 2.5$  and  $\Lambda = 216$  for Data MNTASUB

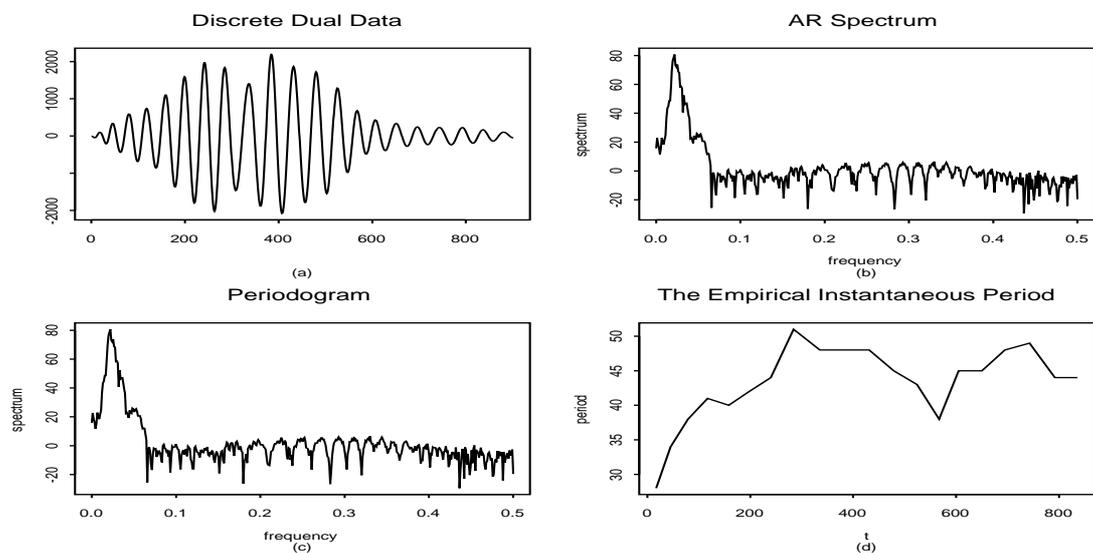


Figure 30: (a) The discrete dual data. (b) AR spectrum of the discrete dual data. (c) Sample spectrum of the discrete dual data. (d) The empirical instantaneous period of the discrete dual data.

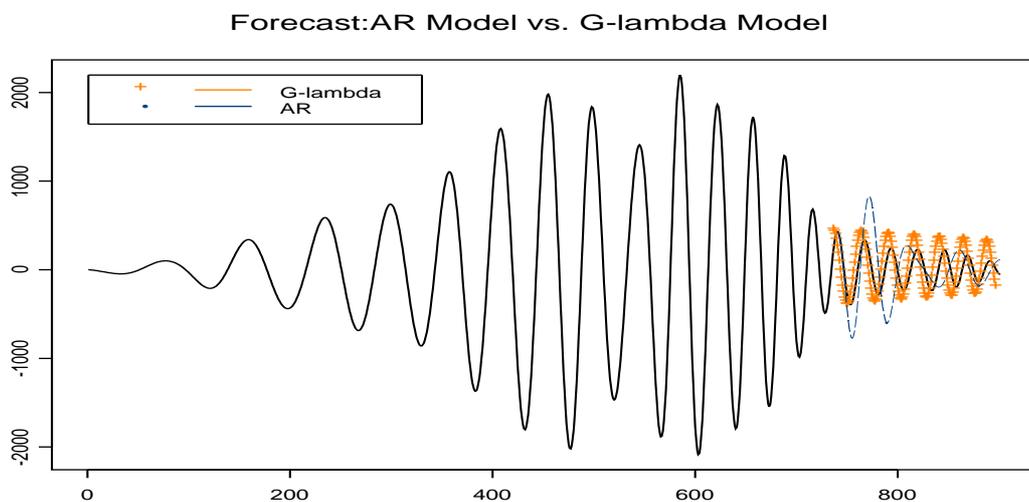


Figure 31: Forecasts for last 120 lags: AR(20) model (dashed line) vs  $G(15,0;2.5)$  model (the line with sign “+”), where the solid line represents the true values.

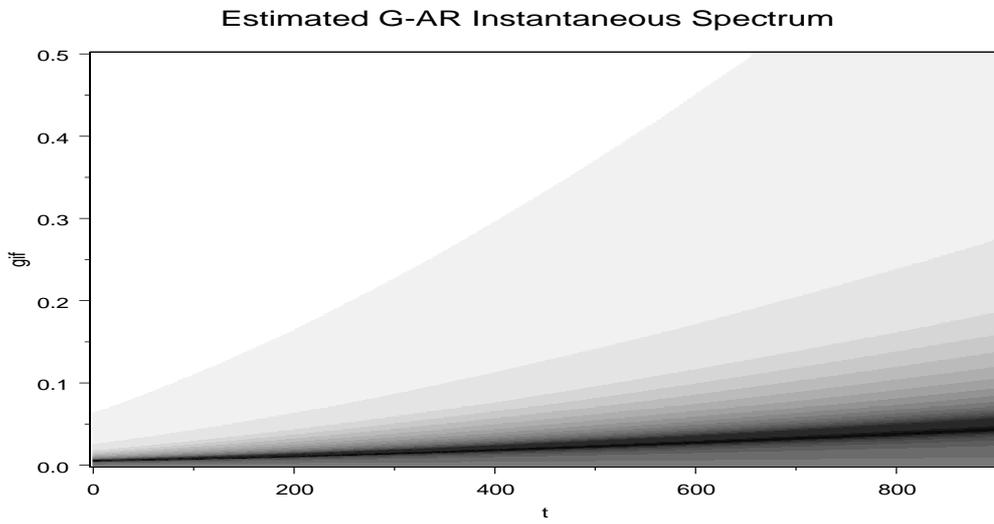


Figure 32: The model-based instantaneous spectrum

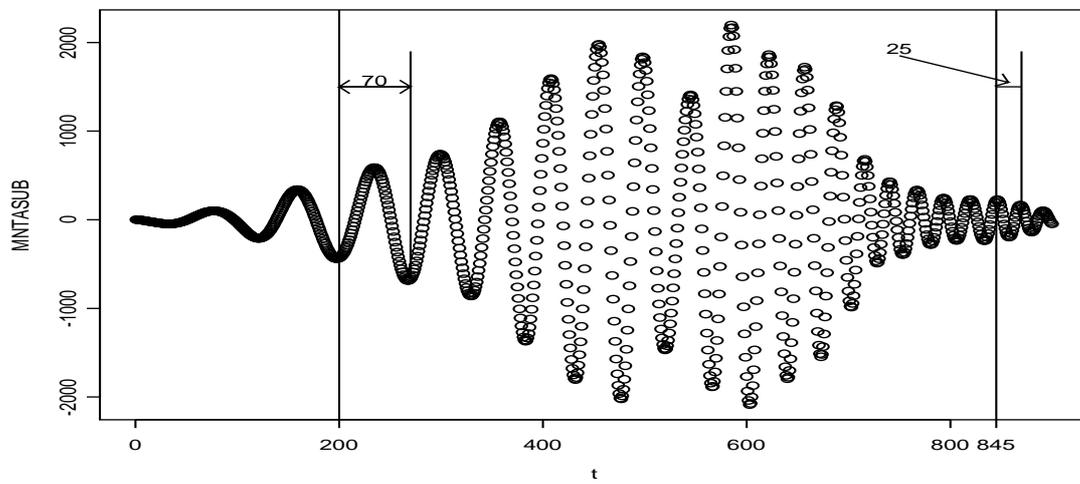


Figure 33: Data: MNTASUB

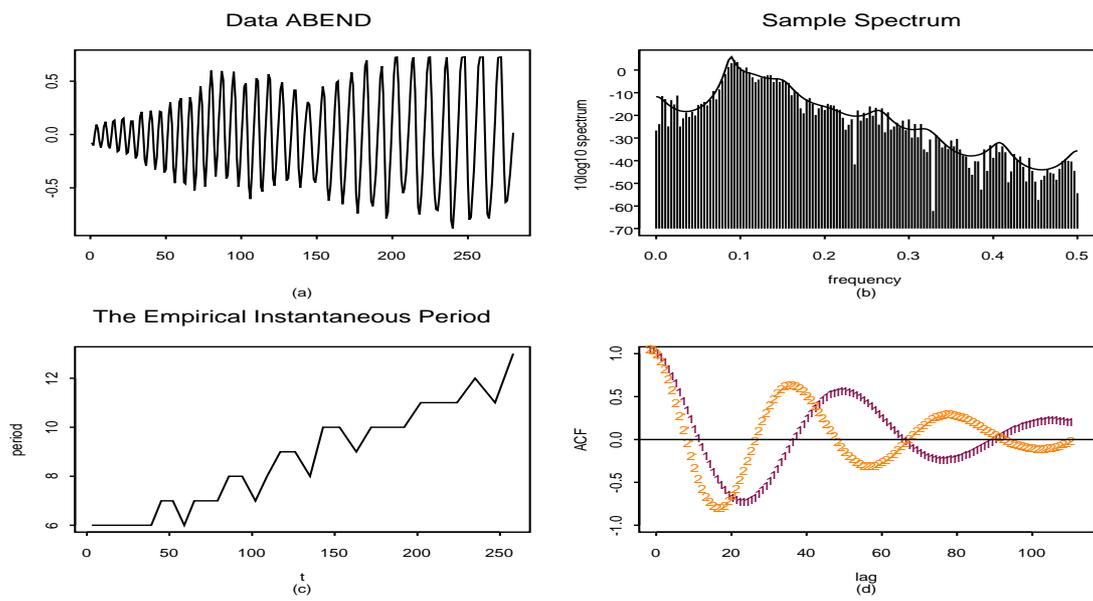


Figure 34: (a) Data ABENDSUB (b)The sample spectrum of Data ABENDSUB. (c) The empirical instantaneous period of Data ABENDSUB. (d)the ACF,  $\hat{\rho}_1$ , (the line with sign “1”) of the first 450 data points vs. the ACF,  $\hat{\rho}_2$ , (the line with sign “2”) of the remaining data points of Data ABENDSUB.

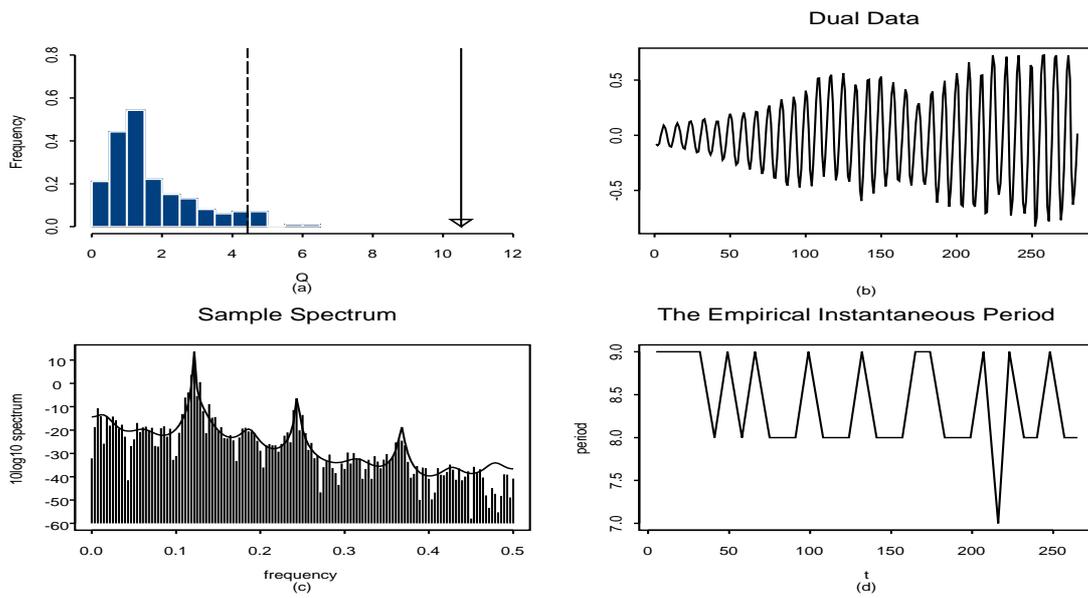


Figure 35: (a) The empirical distribution of  $Q$ ,  $Q_{.95}$ (vertically dashed line) and the observation value(the line with arrow) of  $Q$  for Data ABENDSUB.(b) the dual data. (c) The sample spectrum of the dual data. (d) the empirical instantaneous period of the dual

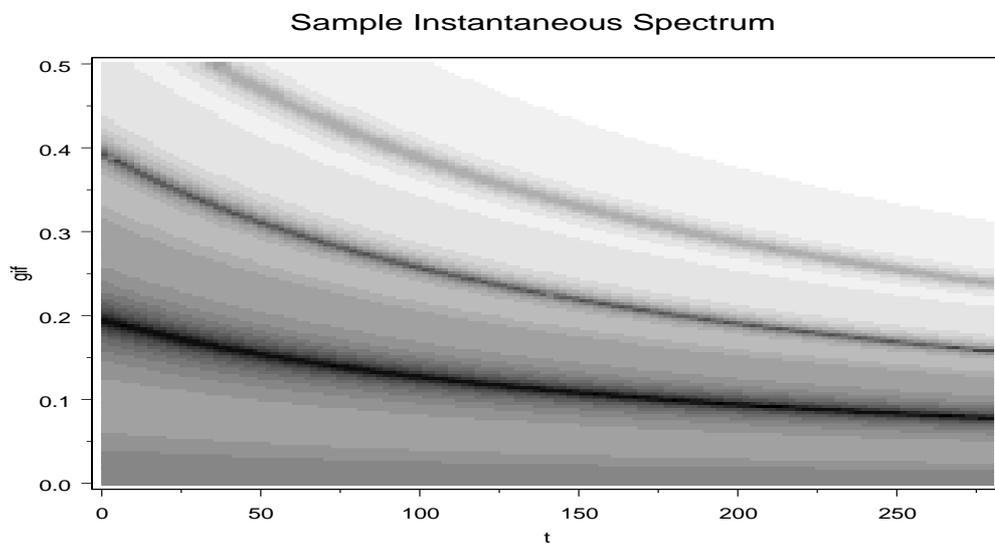


Figure 36: The sample instantaneous spectrum of Data ABENDSUB given  $\lambda = 0$  and  $\Lambda = 188$

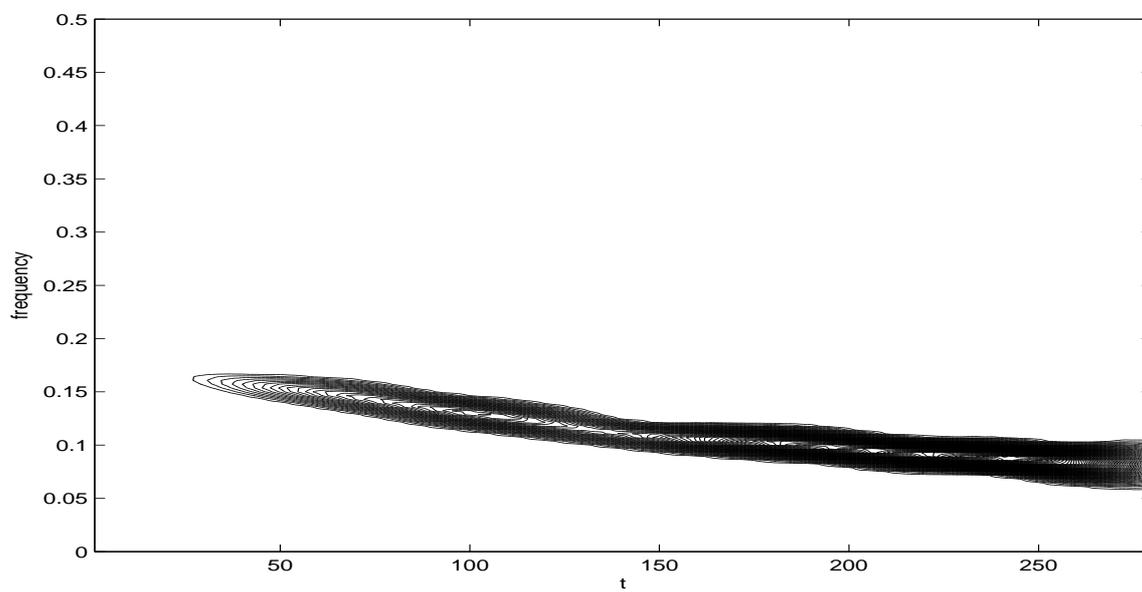


Figure 37: Window-based Fourier analysis for Data ABENDSUB.

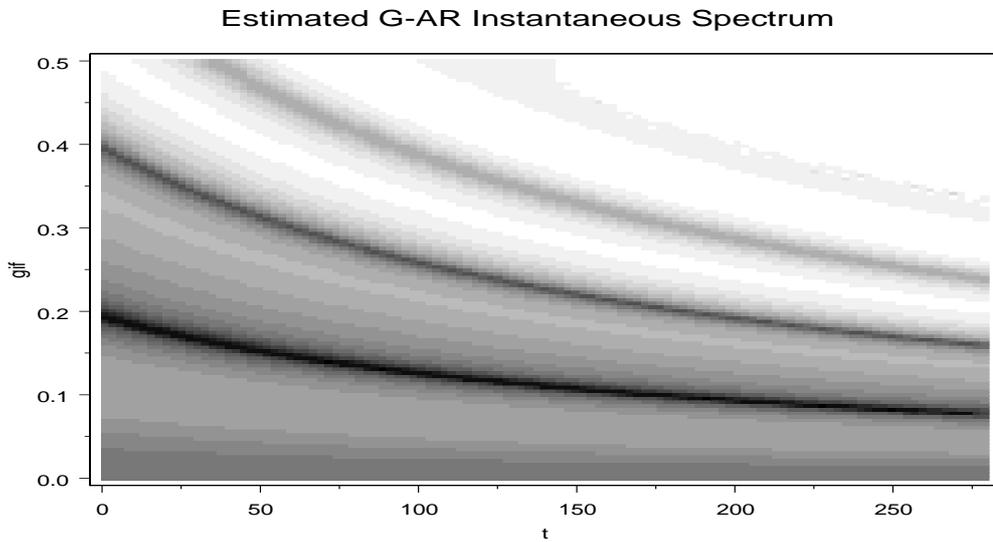


Figure 38: The estimated G-AR instantaneous spectrum of Data ABENDSUB given Euler(11) model.

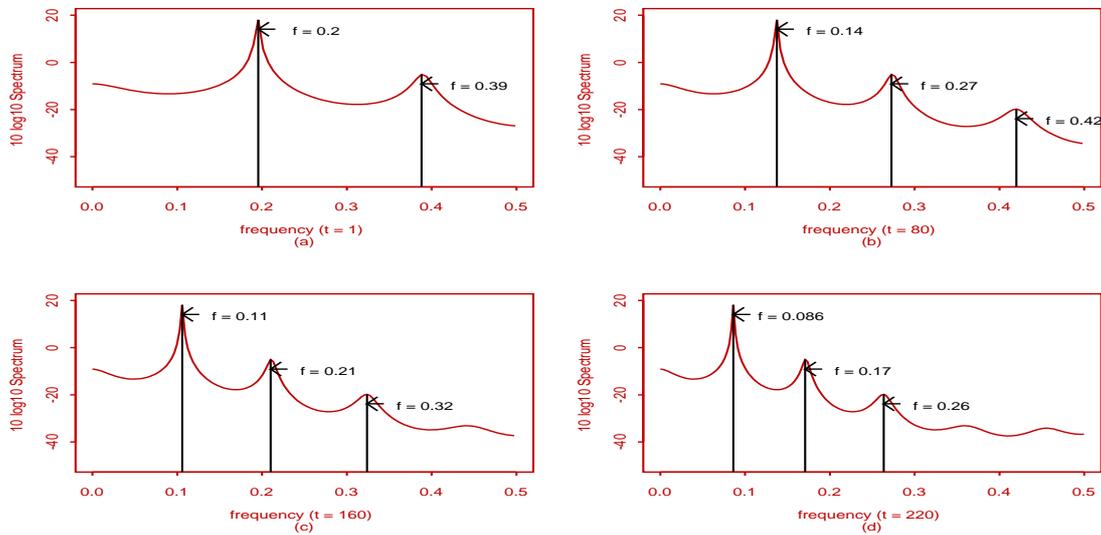


Figure 39: Model-based instantaneous spectrum for Data ABENDSUB at (a)  $t=1$ , (b)  $t=80$ , (c)  $t=160$  and (d)  $t=220$ .