Naive Business Cycle Theory

by

Thomas J. Sargent

Revised March 12, 1975

Working Paper #23

Rsch. File #252.1

Naive Business Cycle Theory

Deterministic (nonrandom) difference operations of low order can generate "cycles," but not of the kind ordinarily thought to characterize economic variables. For example, we have seen that second order difference equations can generate cycles of constant periodicity that are damped, explosive, or, in the very special case where the amplitude $r = 1$, of constant-amplitude. But the "cycles" in economic variables seem neither damped nor explosive, and they don't have a constant period from one cycle to the other; e.g., some recessions last one year, some last for one and a half years. The "business cycle" is the tendency of certain economic variables to possess persistent cycles of approximately constant amplitude and somewhat irregular periodicity from one "cycle" to the other. The National Bureau of Economic Research has inspected masses of data indicating the presence of a business cycle of average length of about three years from peak to peak in many important economic aggregates for the U.S.

Figure 1 graphs the 91 day Treasury Bill rate and the unemployment rate over the postwar period for quarterly data. The "business cycle" shows up in both series, interest rates tending to be high and unemployment low in "booms," and interest rates tending to be low and unemployment high in recessions. Clearly the "cycles" are irregular in length and don't "look like" those generated by our low order difference equations.

While low order deterministic difference equations don't provide an adequate model for explaining the cycles in economic data, low order stochastic or random difference equations do. If the initial condition of a deterministic difference equation is subjected to repeated random shocks of a certain kind, there emerges the possibility of persistent cycles of the kind seemingly infesting economic data. This is an important idea in macroeconomics, and owes its origin to Slutsky and Frisch. These notes sketch the elements of that idea.

A basic building block is the serially uncorrelated random process ε_t, which satisfies

1a $\quad E(\varepsilon_t) = 0 \qquad$ all t

1b $\quad E(\varepsilon_t^2) = \sigma_\varepsilon^2 \qquad$ all t

1c $\quad \text{cov}\,(\varepsilon_t, \varepsilon_{t-s}) = E(\varepsilon_t \varepsilon_{t-s}) = 0 \qquad$ all t and all $s \neq 0$

where E is the mathematical expectations operator. According to (1), the mean of ε_t, which is zero for all t, and the variance of ε_t, which is σ_ε^2 for all t, both are independent of time. According to (1c), ε_t is uncorrelated (i.e., has zero covariance) with itself lagged $s = \pm 1$, $\pm 2, \ldots$ times and is said to be "serially uncorrelated." The variate ε_t is also said to be a "white noise." The schedule of covariances $E(\varepsilon_t \varepsilon_{t-s})$ for $s = 0, \pm 1, \pm 2, \ldots$ is a function only of s, and not of t, a characteristic called "covariance stationarity." The schedule of covariances $E(\varepsilon_t \varepsilon_{t-s})$ is called the "covariogram" of the ε process. Notice that the covariogram, viewed as a function of s, is symmetric about zero since $E(\varepsilon_t \varepsilon_{t-s}) = E(\varepsilon_t \varepsilon_{t+s})$, an implication of $E(\varepsilon_t \varepsilon_{t-s})$ depending only on s and not on t.

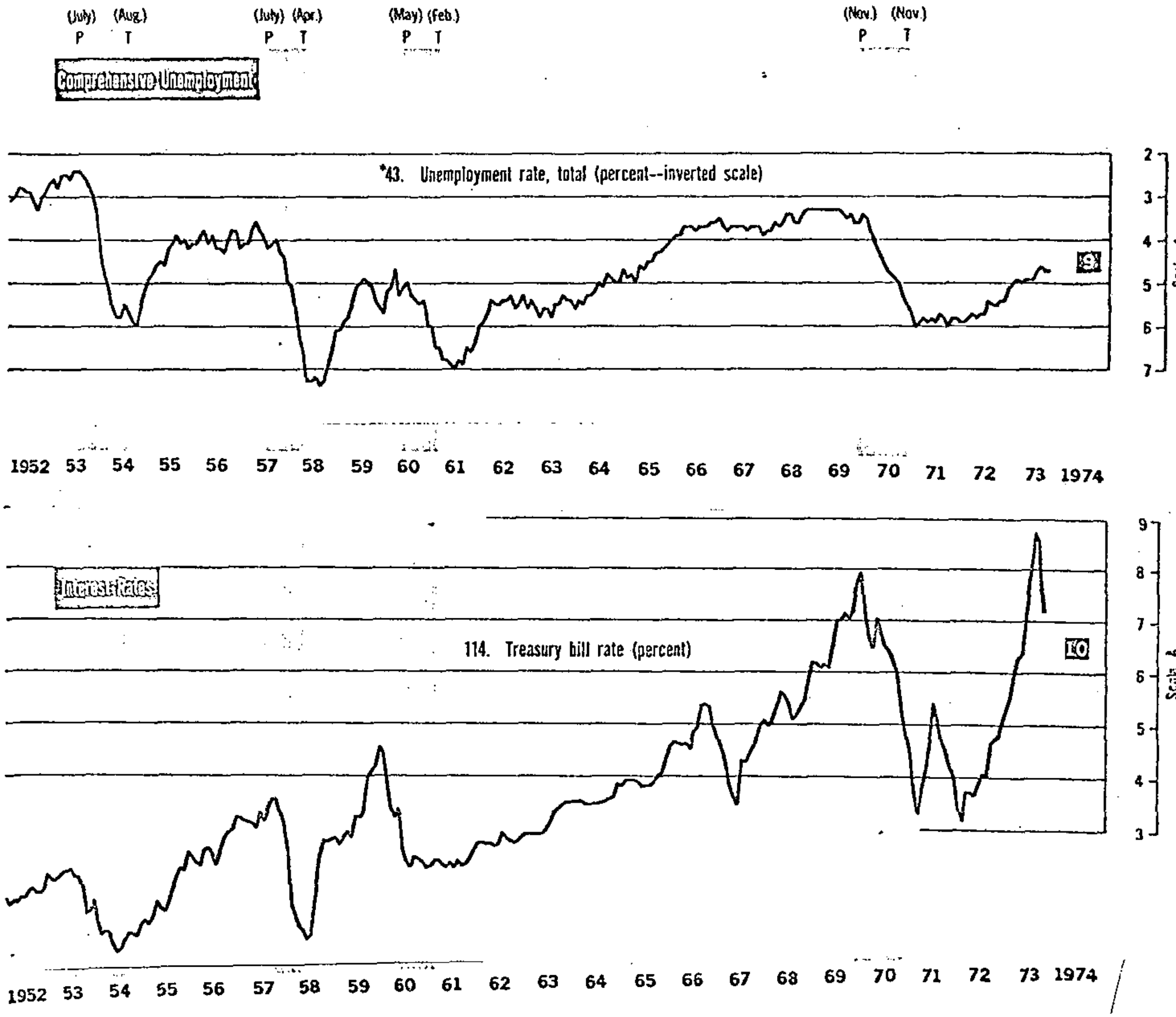
Now consider the random process y_t defined by

$$\text{(a)} \qquad y_t = \sum_{j=0}^{\infty} b_j \, \varepsilon_{t-j}$$

$$= B(L)\varepsilon_t$$

2a

FIGURE 1.


(July) (Aug.) (July) (Apr.) (May) (Feb.) (Nov.) (Nov.)
P T P T P T P T
Comprehensive Unemployment
*43. Unemployment rate, total (percent--inverted scale)
2 3 4 5 6 7
1952 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 1974
Interest Rates
114. Treasury bill rate (percent)
9 8 7 6 5 4 3
1952 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 1974

where $B(L) = \sum_{j=0}^{\infty} b_j L^j$, and where we assume $\sum_{j=0}^{\infty} b_j^{\,2} < \infty$,

a requirement needed to assure that the variance of y is finite. We assume that the ε process is "white" and thus satisfies properties (1). Equation (2) says that the y process is a one-sided moving sum of a white noise process, ε.

We seek the covariogram of the y process, i.e., we seek the values of $c_y(k) = E(y_t y_{t-k})$ for all k. It will be convenient to obtain the "covariance generating function" $g_y(z)$ which is defined by

$$(3) \qquad g_y(z) = \sum_{k=-\infty}^{\infty} c_y(k) z^k.$$

The coefficient on z^k in (3) is the k^{th} lagged covariance, $c_y(k)$.

First notice that taking mathematical expectations on both sides of (2) gives

$$E(y_t) = \sum_{j=0}^{\infty} b_j E(\varepsilon_{t-j})$$

$$= 0 \qquad \text{for all } t.$$

It therefore follows that

$$c_y(k) = E\{(y_t - Ey_t)(y_{t-k} - Ey_{t-k})\}$$

$$= Ey_t \cdot y_{t-k} \qquad \text{for all } k.$$

Notice $y_t \cdot y_{t-k}$ is

$$y_t y_{t-k} = \sum_{j=0}^{\infty} b_j \, \varepsilon_{t-j} \sum_{h=0}^{\infty} b_h \varepsilon_{t-k-h}$$

$$= (b_0\varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + \ldots)(b_0\varepsilon_{t-k} + b_1\varepsilon_{t-k-1} + b_2\varepsilon_{t-k-2} + \ldots)$$

$$y_r y_{t-k} = \{b_0 b_k \varepsilon_{t-k}^2 + b_1 b_{k+1} \varepsilon_{t-k-1}^2 + b_2 b_{k+2} \varepsilon_{t-k-2}^2 + \dots\} +$$

crossproduct terms whose expectations are zero.

Thus

(4) $$c_y(k) = Ey_t y_{t-k} = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} b_j b_{j+k}.$$

The covariance generating function is then

$$g_y(z) = \sum_{k=-\infty}^{\infty} z^k c_y(k)$$

$$= \sigma_\varepsilon^2 \sum_{k=-\infty}^{\infty} z^k \sum_{j=0}^{\infty} b_j b_{j+k}$$

$$= \sigma_\varepsilon^2 \sum_{k=-\infty}^{\infty} \sum_{j=0}^{\infty} b_j b_{j+k} z^k$$

$$g_y(z) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} b_j b_{j+k} z^k.$$

Let $h = j + k$, so that $k = h - j$. Writing the above line in terms of the index h then gives

$$g_y(z) = \sigma_\varepsilon^2 \sum_{j=0}^{\infty} \sum_{h=0}^{\infty} b_j b_h \, z^{h-j}$$

$$= \sigma_\varepsilon^2 \sum_{j=0}^{\infty} b_j z^{-j} \sum_{h=0}^{\infty} b_h z^h.$$

The last equation gives the convenient expression

(5) $$g_t(z) = \sigma_\varepsilon^2 B(z^{-1}) B(z)$$

where $B(z^{-1}) = \sum_{j=0}^{\infty} b_j z^{-j}$, $B(z) = \sum_{j=0}^{\infty} b_j z^j$.

Equation (5) gives the covariance generating function $g_y(z)$ in terms of the b_j's of (2) and the variance σ_ε^2 of the white noise ε.

To take an example that illustrates the usefulness of (5), consider the first order process

$$(6) \qquad y_t = \left(\frac{1}{1-\lambda L}\right)\varepsilon_t = \sum_{i=0}^{\infty} \lambda^i \varepsilon_{t-i}, \qquad |\lambda| < 1$$

where, as always, ε is a white noise process with variance σ_ε^2. We have

$$B(L) = \frac{1}{1-\lambda L},$$

$$B(z) = \frac{1}{1-\lambda z} = 1 + \lambda z + \lambda^2 z^2 + \ldots$$

$$B(z^{-1}) = \frac{1}{1-\lambda z^{-1}} = 1 + \lambda z^{-1} + \lambda^2 z^{-2} + \ldots.$$

(Thus, $B(z)$ is found by replacing L in $B(L)$ by z.) So applying (5), we have

$$(7) \qquad g_y(z) = \sigma_\varepsilon^2 \left(\frac{1}{1-\lambda z^{-1}}\right)\left(\frac{1}{1-\lambda z}\right).$$

From our experience with difference equations we know that the expression (7) can be written as a sum

$$(8) \qquad g_y(z) = \frac{k_1 \sigma_\varepsilon^2}{1-\lambda z} + \frac{k_2 \sigma_\varepsilon^2 z^{-1}}{1-\lambda z^{-1}}$$

where k_1 and k_2 are certain constants. To find out what the constants must be, notice that (8) implies

$$g_y(z) = \sigma_\varepsilon^2 k_1 (1+\lambda z+\lambda^2 z^2+\ldots)$$

$$+ \sigma_\varepsilon^2 k_2 (z^{-1}+\lambda z^{-2}+\lambda^2 z^{-3}+\ldots),$$

so that $c_y(0) = k_1 \sigma_\varepsilon^2$ and $c_y(1) = \sigma_\varepsilon^2 \lambda k_1 = \sigma_\varepsilon^2 k_2$.

By direct computation using (6) we note that

$$Ey_t^2 = \sum_{i=0}^{\infty} \lambda^{i2} E\varepsilon_t^2 = \frac{\sigma_\varepsilon^2}{1 - \lambda^2}$$

$$Ey_t y_{t-1} = E\sum_{i=0}^{\infty} \lambda^i \varepsilon_{t-i} \sum_{i=1}^{\infty} \lambda^{i-1} \varepsilon_{t-i} = E\sum_{i=1}^{\infty} \lambda^i \lambda^{i-1} \varepsilon_{t-i}^2$$

$$= \sigma_\varepsilon^2 \lambda \sum_{i=1}^{\infty} \lambda^{(i-1),2} = \frac{\sigma_\varepsilon^2}{1 - \lambda^2} .$$

So for (8) to be correct, we require that

$$k_1 = \frac{1}{1 - \lambda^2}$$

$$k_2 = \frac{\lambda}{1 - \lambda^2} .$$

With these values of k_1 and k_2, we can verify directly that

$$\sigma_\varepsilon^2 \left(\frac{\frac{1}{1-\lambda^2}}{1-\lambda z} + \frac{z^{-1} \left(\frac{\lambda}{1-\lambda^2}\right)}{1-\lambda z^{-1}} \right)$$

$$= \sigma_\varepsilon^2 \cdot \frac{1}{1-\lambda^2} \left[\frac{(1-\lambda z^{-1}) + \lambda z^{-1} - \lambda^2}{(1-\lambda z)(1-\lambda z^{-1})} \right]$$

$$= \sigma_\varepsilon^2 \frac{1}{(1-\lambda z)(1-\lambda z^{-1})} ,$$

so that (8) and (7) are equivalent.

Expression (8) is the more convenient of the two expressions since it yields quite directly,

$$g_y(z) = \sigma_\varepsilon^2 \frac{1}{1-\lambda^2} \left[\frac{1}{1-\lambda z} - \frac{\lambda z^{-1}}{1-\lambda z^{-1}} \right]$$

(9) $$= \sigma_\varepsilon^2 \frac{1}{1-\lambda^2} [\{1+\lambda z+\lambda^2 z^2+\ldots\} + \{\lambda z^{-1}+\lambda^2 z^{-2}+\lambda^3 z^{-3}+\ldots\}].$$

Thus, we have that for the "geometric" process (6),

$$c_y(k) = \frac{\sigma_\varepsilon 2}{1-\lambda^2} \cdot \lambda^{|k|} \qquad k=0,\pm1,\pm2,\ldots$$

The covariance declines geometrically with increases in $|k|$. We require $|\lambda|<1$ in order that the y process have a finite variance.

To get this result more directly multiply y_t by y_{t-k}, $k > 0$, to obtain

$$y_t y_{t-k} = \lambda y_{t-1} y_{t-k} + \varepsilon_t y_{t-k}.$$

Taking expected values on both sides and noting that $E\varepsilon_t y_{t-k} = 0$ gives

$$E(y_t y_{t-k}) = \lambda E(y_{t-1} y_{t-k})$$

or

$$c_y(k) = \lambda c_y(k-1) \qquad k > 0$$

which implies

$$c_y(k) = \lambda^k c_y(0) \qquad k > 0$$

As a second example, consider the second-order process

(10) $$y_t = (\frac{1}{1-\lambda_1 L})(\frac{1}{1-\lambda_2 L})\varepsilon_t, \qquad |\lambda_1+\lambda_2| < 1$$

where ε_t is white noise with variance σ_ε^2.

For (10) we have

$$B(L) = (\frac{1}{1-\lambda_1 L}) \quad (\frac{1}{1-\lambda_2 L})$$

$$B(z) = \left(\frac{1}{1-\lambda_1 z} \right) \left(\frac{1}{1-\lambda_2 z} \right)$$

$$B(z^{-1}) = \left(\frac{1}{1-\lambda_1 z^{-1}} \right) \left(\frac{1}{1-\lambda_2 z^{-1}} \right).$$

Applying formula (5), we have that the covariance generating function is

$$(11) \qquad g_y(z) = \sigma_\varepsilon^2 \frac{1}{(1-\lambda_1 z)} \frac{1}{(1-\lambda_2 z)} \frac{1}{(1-\lambda_1 z^{-1})} \frac{1}{(1-\lambda_2 z^{-1})} .$$

Notice that (10) can be written

$$y_t = \frac{\lambda_1}{\lambda_1-\lambda_2} \left(\frac{1}{1-\lambda_1 L} \right) \varepsilon_t - \left(\frac{\lambda_2}{\lambda_1-\lambda_2} \right) \left(\frac{1}{1-\lambda_2 L} \right) \varepsilon_t$$

$$(12) \qquad y_t = \frac{\lambda_1}{\lambda_1-\lambda_2} \sum_{i=0}^{\infty} \lambda_1^i \varepsilon_{t-i} - \frac{\lambda_2}{\lambda_1-\lambda_2} \sum_{i=0}^{\infty} \lambda_2^i \varepsilon_{t-i} .$$

For y_{t-k}, $k \geq 0$, we have

$$(13) \qquad y_{t-k} = \frac{\lambda_1}{\lambda_1-\lambda_2} \sum_{i=k}^{\infty} \lambda_1^{i-k} \varepsilon_{t-i} - \frac{\lambda_2}{\lambda_1-\lambda_2} \sum_{i=k}^{\infty} \lambda_2^{i-k} \varepsilon_{t-i} .$$

Multiplying (12) and (13) together and taking expectations gives

$$E(y_t y_{t-k}) = \sigma_\varepsilon^2 \left\{ \frac{\lambda_1^2}{(\lambda_1-\lambda_2)^2} \sum_{i=0}^{\infty} \lambda_1^{k+i} \lambda_1^{i} + \frac{\lambda_2^2}{(\lambda_1-\lambda_2)^2} \sum_{i=0}^{\infty} \lambda_2^{k+i} \lambda_2^{i} \right.$$

$$\left. - \frac{\lambda_1\lambda_2}{(\lambda_1-\lambda_2)^2} \sum_{i=0}^{\infty} \lambda_1^{k+i} \lambda_2^{i} - \frac{\lambda_1\lambda_2}{(\lambda_1-\lambda_2)^2} \sum_{i=0}^{\infty} \lambda_2^{k+i} \lambda_1^{i} \right\}$$

$$(14) \qquad E y_t y_{t-k} = \left(\frac{1}{\lambda_1-\lambda_2} \right)^2 \sigma_\varepsilon^2 \left[\frac{\lambda_1^{2+k}}{(1-\lambda_1^2)} + \frac{\lambda_2^{2+k}}{(1-\lambda_2^2)} - \frac{\lambda_1\lambda_2}{1-\lambda_1\lambda_2} (\lambda_1^k+\lambda_2^k) \right]$$

$k \geq 0.$

So (14) and the symmetry of $g_y(z)$ suggests that the appropriate factorization of (11) is

$$(15) \qquad g_y(z) = \left(\frac{1}{\lambda_1-\lambda_2}\right)^2 \sigma_\varepsilon^2 \{ \left(\frac{\lambda_1^2}{(1-\lambda_1^2)} - \frac{\lambda_1\lambda_2}{1-\lambda_1\lambda_2}\right) \left(\frac{1}{1-\lambda_1 z} + \frac{\lambda_1 z^{-1}}{1-\lambda_1 z^{-1}}\right)$$

$$+ \left(\frac{\lambda_2^2}{1-\lambda_2^2} - \frac{\lambda_1\lambda_2}{1-\lambda_1\lambda_2}\right) \left(\frac{1}{1-\lambda_2 z} + \frac{\lambda_2 z^{-1}}{1-\lambda_2 z^{-1}}\right) \} .$$

According to (14) and (15) the covariogram of a y process governed by the second-order process (10) consists of a weighted sum of two geometric decay processes, the decay parameters being λ_1 and λ_2, the inverse roots of the polynomial $(1-\lambda_1 L)(1-\lambda_2 L)$. Expression (14) implies that the covariogram displays damped oscillations if the roots λ_1 and λ_2 are complex conjugates. This can be shown by substituting $\lambda_1 = re^{-iw}$ and $\lambda_2 = re^{iw}$ into (14), and proceeding to analyze (14) as we above analyzed the solution of the deterministic (nonrandom) second order difference equation. An alternative way to reach the same conclusion is as follows. Multiply both sides of (10) by $(1-\lambda_1 L)(1-\lambda_2 L)$ to get

$$(16) \qquad y_t = t_1 y_{t-1} + t_2 y_{t-2} + \varepsilon_t$$

where $t_1 = (\lambda_1+\lambda_2)$ and $t_2 = -\lambda_1\lambda_2$. Multiply (16) by y_{t-k} for $k \geq 0$ to get

$$y_t y_{t-k} = t_1 y_{t-1} y_{t-k} + t_2 y_{t-2} y_{t-k} + \varepsilon_t y_{t-k}.$$

Since $E\, \varepsilon_t y_{t-k} = 0$, we have

$$E(y_t y_{t-k}) = t_1 E(y_{t-1} y_{t-k}) + t_2 (y_{t-2} y_{t-k}) \qquad k \geq 0$$

which shows that $c_y(k)$ obeys the difference equation

$$(17) \qquad c_y(k) = t_1 c_y(k-1) + t_2 c_y(k-2).$$

So the covariogram of a second (n^{th}) order process obeys the solution to the deterministic second (n^{th}) order difference equation examined above. In particular, corresponding to (17) we consider the polynomial

$$(18) \qquad 1 - t_1 k - t_2 k^2 = 0,$$

which has roots $1/\lambda_1$ and $1/\lambda_2$. (We know that $1-t_1k-t_2k$ equals $(1-\lambda_1 k)(1-\lambda_2 k)$, with roots $1/\lambda_1$ and $1/\lambda_2$.) Alternatively, multiply (18) by k^{-2} to obtain

$$k^{-2} - t_1 k^{-1} - t_2 = 0$$

$$(19) \qquad x^2 - t_1 x - t_2 = 0 \quad \text{where } x = k^{-1}.$$

Notice that the roots of (19) are the reciprocals of the roots of (18), so λ_1 and λ_2 are the roots of (19).

The solution to the deterministic difference equation (17) is, as we have seen,

$$(20) \qquad c_y(k) = \lambda_1^k z_0 + \lambda_2^k z_1, \qquad k \geq 0$$

where z_0 and z_1 are certain constants chosen to make $c_y(0)$ and $c_y(1)$ equal the proper quantities. If the roots λ_1 and λ_2 are complex, we know from our work with deterministic difference equations that (20) becomes

(21) $$c_y(k) = z_0 \frac{r^k}{\sin w} \sin wk + z_1 \frac{r^k}{\sin w} \cos wk$$

where $\lambda_1 = re^{iw}$ and $\lambda_2 = re^{-iw}$. Accordingly to (21), the covariogram displays damped (we require $r < 1$) oscillations with angular frequency w. A complete cycle occurs as wk goes from zero ($k=0$) to 2π ($k=2\pi/w$, if that is possible). So the cycles in the covariogram occur with period from peak to peak of $2\pi/w$ periods. The restrictions on t_1 and t_2 needed to deliver complex roots and so an oscillatory covariogram can be read directly from Figure 1 of "Notes on Difference Equations."

Figure 1 displays some realizations of second order processes for various values of t_1 and t_2, values for which the roots are complex. Notice the tendency of these series to cycle, but with a periodicity that is somewhat variable from cycle to cycle.

The foregoing suggests one definition of the business cycle: a series may be said to possess a "cycle" if its covariogram is characterized by oscillations. The typical "length" of the cycle can be measured by the number of periods it takes for the covariogram to experience one full cycle. To be labelled a business cycle the cycle should exceed a year in length. (Cycles of one year in length are termed "seasonals.")

The Spectrum

An alternative to the preceding definition of the business cycle is based on the spectrum of the y process. Recall the covariance generating function of y defined in (3),

$$g_y(z) = \sum_{k=-\infty}^{\infty} c_y(k) z^k . \tag{3}$$

For the process $y_t = B(L)\varepsilon_t$ we have seen that

$$g_y(z) = B(z)B(z^{-1})\, \sigma_\varepsilon^{\ 2} .$$

If we evaluate (3) at the value $z = e^{-iw}$, we have

$$g_y(e^{-iw}) = \sum_{k=-\infty}^{\infty} c_y(k) e^{-iwk} \qquad -\pi < \ w \ < \pi . \tag{22}$$

Viewed as a function of angular frequency w, $g_y(e^{-iw})$ is called the <u>spectrum</u> of y.

The spectrum $g_y(e^{-iw})$ is itself a covariance generating function, which is hardly surprising. Given an expression for $g_y(e^{iw})$, it is easy to recover the covariances $c_y(k)$. To see this, we multiply (22) by e^{iwh} and integrate with respect to w from $-\pi$ to π:

$$\int_{-\pi}^{\pi} g_y(e^{-iw}) e^{iwh} dw = \int_{-\pi}^{\pi} \sum_{k=-\infty}^{\infty} c_y(k) e^{iw(h-k)} dw \tag{23}$$

$$= \sum_{k=-\infty}^{\infty} c_y(k) \int_{-\pi}^{\pi} e^{iw(h-k)} dw.$$

Now for h = k we have

$$\int_{-\pi}^{\pi} e^{iw(h-k)} dw = \int_{-\pi}^{\pi} 1 dw = 2\,\pi.$$

For $h \neq k$ we have,

$$\int_{-\pi}^{\pi} e^{iw(h-k)} dw = \int_{-\pi}^{\pi} \cos w\,(h-k) dw + i \int_{-\pi}^{\pi} \sin w\,(h-k)\, dw$$

$$= -\sin w(h-k)\Big]_{-\pi}^{\pi} + i \cos w\,(h-k)\Big]_{-\pi}^{\pi}$$

$$= 0.$$

Therefore (23) becomes

$$\int_{-\pi}^{\pi} g_y(e^{-iw}) e^{iwh} dw = 2\pi c_y(h).$$

Thus multiplying the spectrum by e^{iwh} and integrating from $-\pi$ to π gives the h^{th} lagged covariance times 2π. In particular, notice that for $h = 0$, we have

$$\int_{-\pi}^{\pi} g_y(e^{-iw}) dw = 2\pi c_y(0),$$

so that the area under the spectrum from $-\pi$ to π equals 2π times the variance of y. This fact motivates the interpretation of the spectrum as a device for decomposing the variance of a series by frequency. The portion of the variance of the series occurring between any two frequencies is given by the area under the spectrum between those two frequencies.

Notice that from (22) we have

$$g_y(e^{-iw}) = \sum_{k=-\infty}^{\infty} c_y(k) \quad e^{-iwk}$$

(34) $$= c_y(0) + \sum_{k=1}^{\infty} c_y(k)(e^{iwk}+e^{-iwk})$$

$$= c_y(0) + 2 \sum_{k=1}^{\infty} c_y(k) \cos wk.$$

According to (34) the spectrum is real valued at each frequency, and is obtained by multiplying the covariogram of y by a cosine function of the frequency in question. Notice also that since cos x = cos -x, it follows from (34) that

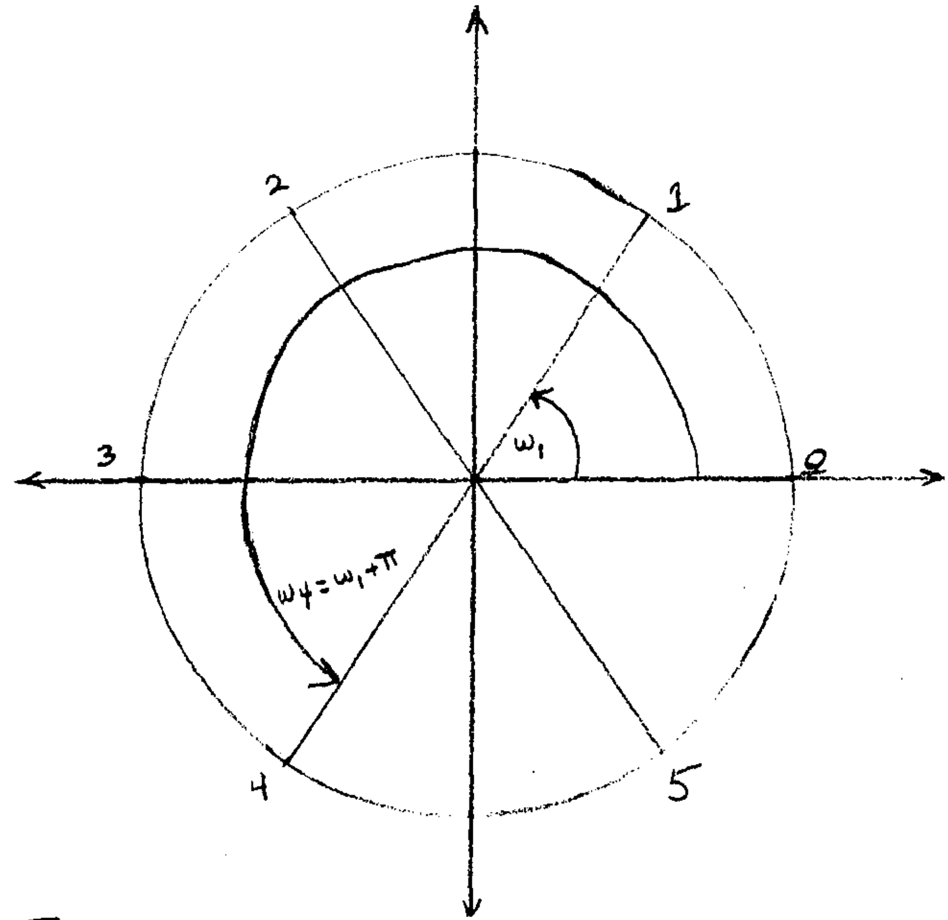
$$g_y(e^{iw}) = g_y(e^{-iw}),$$

so that the spectrum is symmetric about w=0.

Notice also that since $\cos(w+2\pi k) = \cos(w)$, $k=0, \pm 1, \pm 2, \ldots$, it follows that the spectrum is a periodic function of w with period 2π. Therefore we can confine our attention to the interval $[-\pi,\pi]$, or even $[0,\pi]$ by virtue of the symmetry of the spectrum about w=0.

The fact that the spectrum can be viewed as decomposing the variance of a series by frequency motivates our second definition of the business cycle. A series is said to display a cycle of a given periodicity if its spectrum possesses a peak at that periodicity. A series displays a "business cycle" if that periodicity is of about three years. If a peak occurs in a spectrum at a certain frequency, it indicates that a relatively large amount of the variance of the series occurs (can be explained by cosine functions) at that frequency. The sharper is the peak in the spectrum, the more regular are the cycles occurring in the series.

To motivate further the interpretation of the spectrum as a decomposition of variance by frequency, suppose that we have T observations on y_t, $t=0, 1, \ldots, T-1$. Suppose that T is an even number. We consider computing the following regression of y_t on sine and cosine functions of angular frequency $w_j = \frac{2\pi j}{T}$ where $j=0, 1, \ldots, T/2$:



$T = 6$

$$(25) \qquad y_t = \sum_{k=0}^{T/2} \hat{\alpha}(w_k) \cos w_k t + \sum_{k=1}^{T/2-1} \hat{\beta}(w_k) \sin w_k t, \; w_k = \frac{2\pi k}{T} .$$

There are T observations and T independent variables in (25). The independent variables of (25) are mutually orthogonal. For we know that

$$\cos w_k t \cos w_j t + \sin w_k t \sin w_j t = \cos (w_k - w_j) t$$

$$\cos w_k t \cos w_j t - \sin w_k t \sin w_j t = \cos (w_k + w_j) t .$$

Summing both equations and adding we have for $j \neq k$

$$(26) \qquad 2 \sum_{t=0}^{T-1} \cos w_k t \cos w_j t = \sum_{t=0}^{T-1} \cos (w_k - w_j) t + \sum_{t=0}^{T-1} \cos (w_k + w_j) t$$

$$= 0,$$

since the angles $(w_k \pm w_j) t = \frac{2\pi (j \pm k) t}{T}$, $t=0, 1, \ldots, T-1$ are spaced evenly about the circle in the fashion depicted in Figure ____. The angles appear in pairs, w', $w' + \pi$, so that for each cosine in the sum of angle w', there is another offsetting cosine associated with the angle $w' + \pi$. From (26) it follows that

$$\sum_{t=0}^{T-1} \cos w_k t \cos w_j t = 0, \qquad k \neq j$$

so that $\cos w_k t$ and $\cos w_j t$ are orthogonal. In a similar fashion, it can be shown that

$$\sum_{t=0}^{T-1} \sin w_k t \cos w_j t = \sum_{t=0}^{T-1} \sin w_k t \sin w_j t = 0 \text{ for } j \neq k,$$

so that the independent variables are mutually orthogonal.

Where the independent variables are orthogonal, the (multivariate) least squares estimator of the regression coefficients is identical with the vector of simple least squares estimates. These are given by

$$\hat{\alpha}(w_k) = \frac{\sum_{t=0}^{T-1} y_t \cos w_k t}{\sum_{t=0}^{T-1} \cos^2 w_k t} \qquad k=0,1,\ldots T/2$$

(27)

$$\hat{\beta}(w_k) = \frac{\sum_{t=0}^{T-1} y_t \sin w_k t}{\sum_{t=0}^{T-1} \sin^2 w_k t} . \qquad k=1,2,\ldots T/2-1$$

Notice that

$$\sum_{t=0}^{T-1} \cos^2 w_0 t = \sum_{t=0}^{T-1} 1 = T$$

$$\sum_{t=0}^{T-1} \cos^2 w_{T/2} = \sum_{t=0}^{T-1} \cos^2 (\pi t) = T .$$

and for $k = 1, 2, \ldots, T/2 - 1$

$$\sum_{t=0}^{T-1} \cos^2 w_k t = \sum_{t=0}^{T-1} (\cos 2 w_k t + \sin^2 w_k t) = \sum_{t=0}^{T-1} \sin^2 w_k t$$

$$= \sum_{t=0}^{T-1} (1-\sin^2 w_k t)$$

which implies that

$$\sum_{t=0}^{T-1} \sin^2 w_k t = \sum_{t=0}^{T-1} \cos^2 w_k t = T/2 \quad \text{for } k = 1, 2, \ldots, T/2 - 1.$$

Thus, (27) becomes

$$\hat{\alpha}(w_0) = \frac{\sum_{t=0}^{T-1} y_t}{T}$$

(28) $$\hat{\alpha}(w_{T/2}) = \frac{1}{T} \sum_{t=0}^{T-1} y_t (-1)^t$$

$$\hat{\alpha}\ (w_k) = \frac{2}{T} \sum_{t=0}^{T-1} y_t \cos w_k t \qquad k = 1, 2, \ldots, T/2 - 1$$

$$\hat{\beta}\ (w_k) = \frac{2}{T} \sum_{t=0}^{T-1} y_t \sin w_k t \qquad k = 1, 2, \ldots, T/2 - 1.$$

Since (25) represents a regression of T observations on y against T orthogonal independent variables (which guarantees that the X'X matrix of the linear statistical model is of full rank), we know that the regression fits the data exactly, i.e., it gives a perfect fit. So what we have achieved is a decomposition of y_t $\{t=0, \ldots, T-1\}$ into a weighted sum of sine and cosine terms of angular frequencies $w_k = \frac{2\pi k}{T}$, $k = 0, \ldots, T/2$. The least squares regression coefficients $\hat{\alpha}\ (w_k)$ and $\hat{\beta}\ (w_k)$ give a measure of how important the various frequencies are in composing the series y_t. To make this more precise, notice that from (25), the sample variance of the y's can be written

$$\frac{1}{T} \sum_{t=0}^{T-1} \left(y_t - \frac{\sum_{t=0}^{T-1} y_t}{T}\right)^2 = \frac{1}{T} \sum_{t=0}^{T-1} \left(y_t - \hat{\alpha}(w_0)\right)^2$$

$$= \frac{1}{T} \left\{ \sum_{k=1}^{T/2-1} \hat{\alpha}(w_k)^2 \sum_{t=0}^{T-1} \cos^2 (w_k t) + \sum_{k=1}^{T/2-1} \hat{\beta}(w_k)^2 \sum_{t=0}^{T-1} \sin^2 (w_k t) \right.$$

$$\left. + \ \hat{\alpha}(w_{T/2})^2 \sum_{t=0}^{T-1} \cos^2 (w_{T/2} t) \right\},$$

which follows by virtue of the orthogonality of sines and cosines of different frequencies. From our earlier calculations of $\sum_{t=0}^{T-1} \cos^2 w_k t$ and $\sum_{t=0}^{T-1} \sin^2 w_k t$, the above equation becomes

$$\frac{1}{T}\sum_{t=0}^{T-1}(y_t - \frac{\sum y_t}{T})^2 = \frac{1}{T}\{[T/2\sum_{k=1}^{T/2-1} \hat{\alpha}^2(w_k) + \hat{\beta}^2(w_k)]$$

$$+ T\,\hat{\alpha}^2(w_{T/2})\} = \frac{1}{2}\{\sum_{t=1}^{T/2-1}[\hat{\alpha}^2(w_k) + \hat{\beta}^2(w_k)] + 2\,\hat{\alpha}^2(w_{T/2})\}.$$

Thus, the term $1/2\ [\hat{\alpha}^2(w_k) + \hat{\beta}^2(w_k)]$ measures the contribution of sine and cosine terms of frequency w_k to the sample variance of y.

To look at the coefficients $\hat{\alpha}(w_k)$ and $\hat{\beta}(w_k)$ from a slightly different perspective, consider the quantities

$$(29)\qquad A(w_k) = \frac{1}{T}\sum_{t=0}^{T-1} y_t e^{iw_k t} = \frac{1}{T}\sum_{t=0}^{T-1} y_t \cos w_k t + i\,\frac{1}{T}\sum_{t=0}^{T-1} y_t \sin w_k t$$

$$\equiv a(w_k) + ib(w_k) \qquad w_k = \frac{2\pi k}{T},\ k=0,\ 1,\ \ldots,\ T-1$$

where $a(w_k) = \frac{1}{T}\sum_{t=0}^{T-1} y_t \cos w_k t$, $b(w_k) = \frac{1}{T}\sum_{t=0}^{T-1} y_t \sin w_k t$.

The list of $A(w_k)$'s for k=0, 1, ..., T-1 is known as the Fourier transform of the series y_t {t=0, ..., T-1}.

Now consider the quantity

$$(30)\qquad \sum_{k=0}^{T-1} A(w_k)\cdot e^{-iw_k t}$$

$$= \sum_{k=0}^{T-1}(a(w_k) + ib(w_k))(\cos w_k t - i \sin w_k t).$$

Notice that $w_{T-k} = \frac{2\pi(T-k)}{T} = 2\pi - \frac{2\pi k}{T} = 2\pi - w_k$.

Since $\sin(x+2\pi) = \sin(x)$ and $\cos(x+2\pi) = \cos x$, it follows that $A(w_k) = A(w_k+2\pi)$. Furthermore, since $-\sin x = \sin(-x)$ and $\cos(x) = \cos(-x)$, it follows that

$$A(w_{T-k}) = A(2\pi - w_k) = A(-w_k)$$

$$= a(-w_k) + ib(-w_k)$$

$$= a(w_k) - ib(w_k) \equiv \overline{A(w_k)}$$

where $\overline{A(w_k)}$ is the complex conjugate of $A(w_k)$. Consequently, (30) can be written as

$$\sum_{k=0}^{T-1} A(w_k)e^{-iw_k t} = (a(w_{T/2}) + ib(w_{T/2}))(\cos w_{T/2}t - i \sin w_{T/2}t)$$

$$- (a(w_0) + ib(w_0))(\cos w_0 t - i \sin w_0 t)$$

$$+ \sum_{k=0}^{T/2-1} (a(w_k)+ib(w_k))(\cos w_k t - i \sin w_k t)$$

$$+ \sum_{k=0}^{T/2-1} (a(w_k)-ib(w_k))(\cos w_k t + i \sin w_k t)$$

which, since $\sin w_{T/2} = 0$, equals

$$- a(w_0) + a(w_{T/2}) \cos (w_{T/2}t) + \sum_{k=0}^{T/2-1} 2\, a(w_k) \cos w_k t + \sum_{k=0}^{T/2-1} 2\, b(w_k) \sin w_k t. \tag{31}$$

Comparing $a(w_k)$ and $b(w_k)$ with our earlier least squares estimates $\hat{\alpha}(w_k)$ and $\hat{\beta}(w_k)$ we notice that

(32)

$$a(w_0) = \frac{1}{T} \sum_{t=0}^{T-1} y_t = \hat{\alpha}(w_0)$$

$$a(w_k) = \frac{1}{2}\hat{\alpha}(w_k) \qquad k = 1, 2, \ldots, T/2 - 1$$

$$a(w_{T/2}) = \hat{\alpha}\,(w_{T/2})$$

$$b(w_k) = \frac{1}{2}\hat{\beta}\,(w_k) \qquad k = 1, 2, \ldots, T/2 - 1.$$

Consequently (31) equals the least squares regression

$$\sum_{k=0}^{T/2} \hat{\alpha}(w_k) \cos w_k t + \sum_{k=1}^{T/2-1} \hat{\beta}(w_k) \sin w_k t$$

which we know equals y_t by virtue of the perfect fit of (25). We have therefore proved that

$$\sum_{k=0}^{T-1} A(w_k)e^{-iw_k t} = y_t,$$

which is a theorem due to Fourier. The real and imaginary parts of $A(w_k) = a(w_k) + i\, b(w_k)$ are (apart from a scaler for $k = 1, \ldots, T/2-1$) the regression coefficients in (25), as is summarized in (32).

A "natural" measure of the importance of the cosine and sine of the frequency w_k in composing y_t is the squared amplitude of $A(w_k)$, which is

$$A(w_k)\overline{A(w_k)} = |A(w_k)|^2$$

$$= (a(w_k)+ib(w_k))(a(w_k)-ib(w_k))$$

$$= a^2(w_k)+b^2(w_k).$$

The higher is this quantity, the larger is the weight put on the sine and cosine of frequency w_k in (25) in making up y_t. As it turns out, the quantity $|A(w_k)|^2$ can be used to estimate the value of the spectrum of y at frequency w_k.

Consider the quantity

$$T\, A(w_k)\overline{A(w_k)} = \frac{1}{T} \sum_{t=0}^{T-1} y_t e^{iw_k t} \sum_{j=0}^{T-1} y_j e^{-iw_j j}$$

$$= \frac{1}{T} \sum_{t=0}^{T} \sum_{j=0}^{T} y_j y_t e^{iw_k (t-j)}$$

$$= \frac{1}{T} \{ y_0(y_0+y_1e^{-iw_k1}+y_2e^{-iw2}+ \dots +y_{T-1}e^{-iw(T-1)})$$

$$+ y_1(y_0e^{iw1}+y_1+y_2e^{-iw1}+ \dots +y_{T-1}e^{-iw(T-2)})$$

$$+ y_2(y_0e^{iw2}+y_2e^{iw1}+y_2+ \dots +y_{T-1}e^{-iw(T-3)})$$

$$+$$

$$\cdot$$

$$\cdot$$

$$\cdot$$

$$+ y_{T-1}(y_0e^{iw(T-1)}+ye^{iw(T-2)}+ \dots +y_{T-1})$$

$$= \frac{1}{T}\sum_{t=0}^{T-1} y_t^2 + \frac{1}{T}\sum_{j=1}^{T-1}\sum_{t=j}^{T-1} y_t y_{t-j}(e^{iw_kj}+e^{-iw_kj})$$

$$TA(w_k)\overline{A(w_k)} = \frac{1}{T}\sum_{t=0}^{T-1} y_t^2 + \frac{1}{T}\sum_{j=1}^{T-1}\sum_{t=j}^{T-1} y_t y_{t-j}\ 2\cos w_kj\ .$$

We are taking the view that y_t $\{ t = 0, 1, \dots, T-1 \}$ is the realization of a random process, so that it is appropriate to inquire about the expected value of $TA(w_k)\overline{A(w_k)}$, which is a random variable itself, being a function of the random y_t's. Taking expected values on both sides of the above equation gives

$$ETA(w_k)\overline{A(w_k)} = ET(a^2(w_k)+b^2(w_k))$$

$$= c_y(0) + 2 \sum_{j=1}^{T-1} \frac{1}{T} \sum_{t=j}^{T-1} E(y_t y_{t-j}) \cos w_k j$$

(33) $$ETA(w_k)\overline{A(w_k)} = c_y(0) + 2 \sum_{j=1}^{T-1} (1-\frac{j}{T}) c_y(j) \cos w_k j.$$

Recall that the spectrum of y is

$$g_y(e^{-iw_k}) = c_y(0) + 2 \sum_{j=1}^{\infty} c_y(j) \cos w_k j.$$

Now as $T \to \infty$, the term $1 - j/T \to 1$ for fixed j. Thus if $c_y(j)$ approaches zero fast enough as $j \to \infty$, we have that

$$ETA(w_k)\overline{A(w_k)} \to g_y(e^{-iw_k})$$

for <u>all</u> frequencies w_k (not just those for which $w_k = 2\pi k/T$, integer k). (It is possible to show that for any T,

$$ETA(w_k)\overline{A(w_k)} = g_y(e^{-iw_k})$$

for $w_k = 2\pi k/T$, k an integer. See, e.g., Melvin Hinich, "Introduction to Fourier Analysis of Data," Center for Naval Analyses, 1969, p. 22.)

We have thus showed that the variable $(a^2(w_k)+b^2(w_k))$ bears an intimate relation to the spectrum. The quantities $(a^2(w_k)+b^2(w_k))$ are called periodogram ordinates, and a graph of them for various w_k against w_k is known as the periodogram. (In fact computing the periodogram is where one often begins on the way to estimating a spectrum.) It is the relation of the quantities $a(w_k)$ and $b(w_k)$ to the regression (25) of y_t on sines and cosines, on the one hand, and to the spectrum of y on the other hand, that motivates the interpretation of the spectrum as a decomposition of the variance of y by frequency.

Let us examine the spectra of some simple processes. First consider the white noise process

$$y_t = \varepsilon_t$$

ε_t white so that $c_y(0) = \sigma_\varepsilon^2$, $c_y(h) = 0$ for $h \neq 0$.

For this process the covariance generating function is simply

$$g_y(z) = \sigma_\varepsilon^2,$$

so that the spectrum is

$$g_y(e^{-iw}) = \sigma_\varepsilon^2, \qquad -\pi \leq w \leq \pi$$

so that the spectrum is flat, as depicted in Figure 3, and equals σ_ε^2 at each frequency. Notice that

$$\int_{-\pi}^{\pi} g_y(e^{-iw})dw = 2\pi\,\sigma_\varepsilon^2,$$

as expected. So a white noise has a flat spectrum, indicating that all frequencies between $-\pi$ and π are equally important in accounting for its variance.

Next consider the first order process

$$y_t = B(L)\,\varepsilon_t = \frac{1}{1-\lambda L}\;\varepsilon_t. \qquad -1 < \lambda < 1.$$

For this process the covariance generating function is

$$g_y(z) = \left(\frac{1}{1-\lambda z}\right)\left(\frac{1}{1-\lambda z^{-1}}\right)\sigma_\varepsilon^2.$$

Therefore the spectrum is

$$g_y(e^{-iw}) = \left(\frac{1}{1-\lambda e^{-iw}}\right)\left(\frac{1}{1-\lambda e^{iw}}\right) \sigma_\varepsilon^2$$

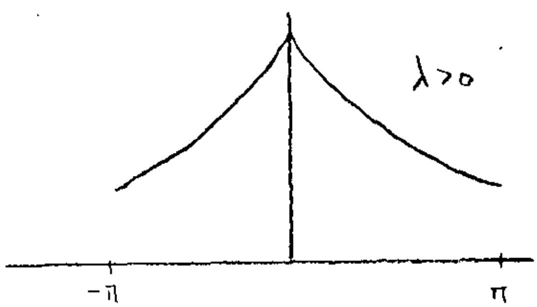
$$= \frac{1}{(1-2\lambda(e^{iw}+e^{-iw}) + \lambda^2)}$$

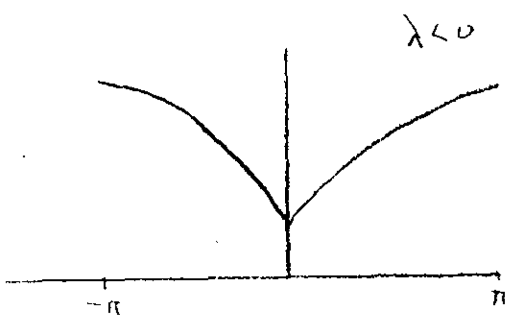
$$g_y(w) = \frac{1}{1-2\lambda \cos w + \lambda^2} .$$

Notice that

$$\frac{dg_y(w)}{dw} = -(1-2\lambda\cos w + \lambda^2)^{-2}(2\lambda \sin w) .$$

The first term in parenthesis is positive. Since $\sin w > 0$ for $0 < w < \pi$, the second term is negative on $(0,\pi]$ if $\lambda < 0$ and positive on $(0,\pi]$ if $\lambda > 0$. Therefore if $\lambda > 0$, the spectrum decreases on $(0,\pi]$ as w increases; if $\lambda < 0$, the spectrum increases on $(0,\pi]$ as w increases.





Thus if $\lambda > 0$, low frequencies (i.e., low values of w) are relatively important in composing the variance of w, while if $\lambda < 0$, high frequencies are the more important. It is easy to verify that the higher in absolute value is λ, the steeper is the spectrum. Notice that the first order process can have a peak in its spectrum only at $w=0$ or $w=\pm\pi$. A peak at $w=\pi$ corresponds to a periodicity of $2\pi/w = 2\pi/\pi = 2$ periods. A peak at $w=0$, corresponds to a cycle with "infinite" periodicity, which is unobservable and hence not a cycle at all.

With quarterly data, a business cycle corresponds to a peak in the spectrum at a periodicity of about 12 quarters. A first order process is capable of having a peak only at two quarters or at "infinite" quarters, and so is not capable of rationalizing a business cycle in the sense of a peak in the spectrum at about 12 quarters. As we saw above, a first order process cannot possess a covariogram with a periodicity other than two periods, and so with quarterly data cannot rationalize a business cycle in the sense of an oscillatory covariogram.

Next consider the second order process

$$y_r = \frac{1}{1-t_1L-t_2L^2} \qquad \varepsilon_t,$$

ε_t white noise. For this process the covariance generating function is

$$g_y(z) = \frac{1}{1-t_1z-t_2z^2} \cdot \frac{1}{1-t_1z^{-1}-t_2z^{-2}} \sigma_\varepsilon^2 .$$

Therefore the spectrum of the process is

$$g_y(e^{-iw}) = \frac{1}{1-t_1e^{-iw}-t_2e^{-2iw}} \quad \frac{1}{1-t_1e^{iw}-t_2e^{2iw}} \quad \sigma_\varepsilon^2$$

$$= \frac{\sigma_\varepsilon^2}{1+t_1^2+t_2^2+(t_2t_1-t_1)(e^{iw}+e^{-iw})-t_2(e^{-2iw}+e^{2iw})}$$

$$= \frac{\sigma_\varepsilon^2}{1+t_1^2+t_2^2-2t_1(1-t_2)\cos w-2t_2\cos 2w} = \frac{\sigma_\varepsilon^2}{h(w)} .$$

Differentiating with respect to w, we have

$$\frac{dg_y(e^{-iw})}{dw} = -\sigma_\varepsilon^2 h(w)^{-2}(2t_1(1-t_2)\sin w + 4t_2 \sin 2w)$$

$$+ -\sigma_\varepsilon^2 h(w)^{-2}(2\sin w \cdot [t_1(1-t_2)+4t_2\cos w]).$$

We know that $h(w) > 0$. For the above derivative to be zero at a w belonging to $(0, \pi)$, we must have the term in brackets equal to zero:

$$t_1(1-t_2) + 4t_2 \cos w = 0$$

or

$$\cos w = \frac{-t_1(1-t_2)}{4t_2} \tag{35}$$

so that

$$w = \cos^{-1}\left(\frac{-t_1(1-t_2)}{4t_2}\right) . \tag{35'}$$

Equation (35) can be satisfied only if

$$\left|\frac{-t_1(1-t_2)}{4t_2}\right| < 1, \tag{36}$$

since $|\cos x| \leq 1$ for all x. If (36) is met, the spectrum of y does achieve a maximum on $(0,\pi)$. Condition (36) is slightly more restrictive than the condition that the roots of the deterministic difference equation be complex so that the covariogram display oscillations. Let us write (36) as

$$-1 < \frac{-t_1(1-t_2)}{4t_2} < 1. \tag{37}$$

The boundaries of the region (37) are

$$-t_1(1-t_2) = 4t_2 \tag{38}$$

and

$$-t_1(1-t_2) = -4t_2 . \tag{39}$$

The points $(t_1, t_2) = (0,0)$ appear on both boundaries, while the point $(t_1, t_2) = (2, -1)$ appears on (38) and $(t_1, t_2) = (-2, -1)$ appears on (39). Differentiating (38) implicitly with respect to t_1 gives

$$\frac{dt_2}{dt_1} = \frac{t_2-1}{4-t_2},$$

so that along (38)

$$\left.\frac{dt_2}{dt_1}\right)_{t_1=t_2=0} = -\frac{1}{4}$$

and

$$\left.\frac{dt_2}{dt_1}\right)_{\substack{t_1=2\\ t_2=-1}} = -1.$$

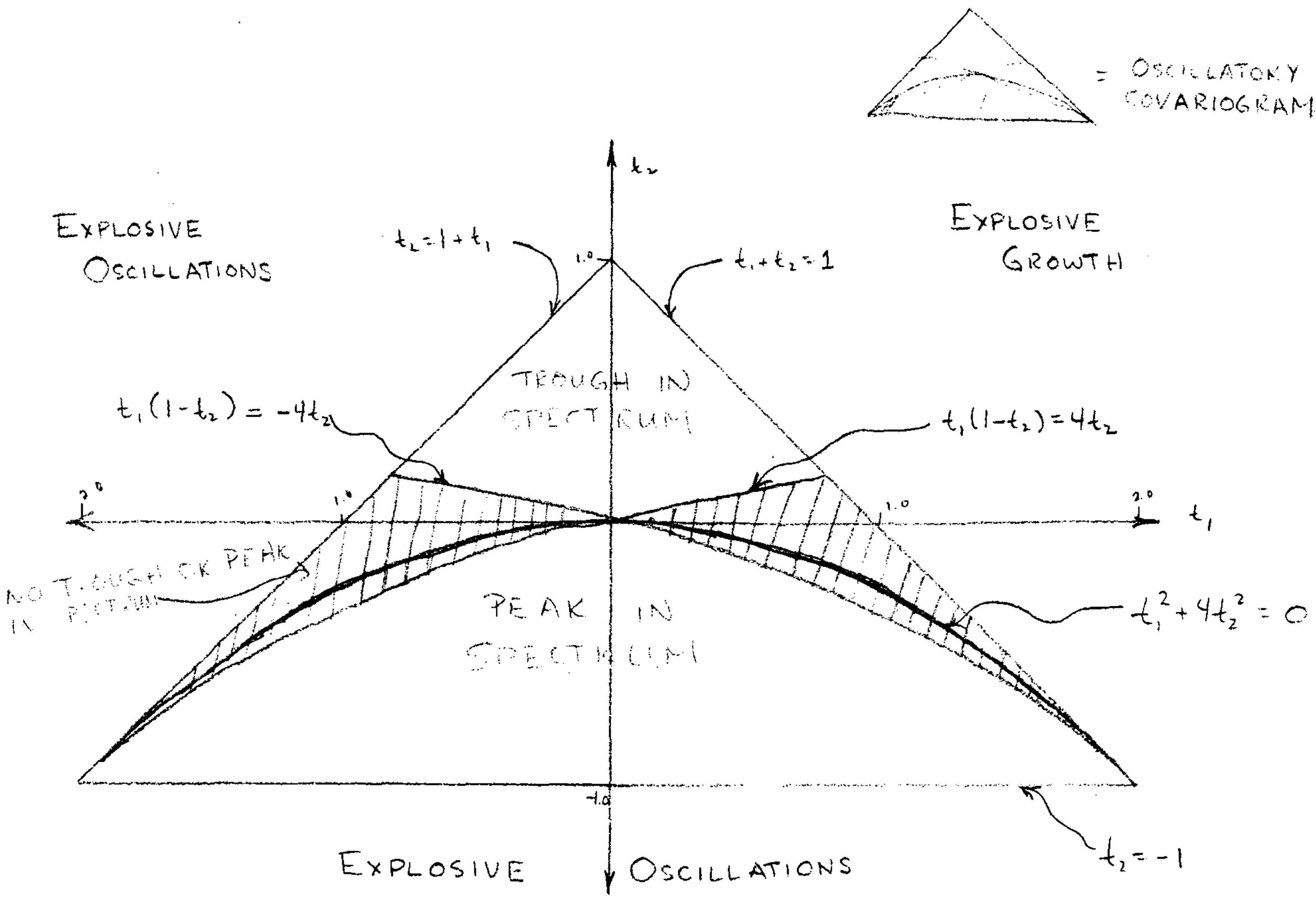
Differentiating (39) with respect to t_1 gives

$$\frac{dt_2}{dt_1} = \frac{1-t_1}{4+t_1}$$

so that along (39)

$$\left.\frac{dt_2}{dt_1}\right)_{t_1=t_2=0} = \frac{1}{4}$$

$$\left.\frac{dt_2}{dt_1}\right)_{\substack{t_1=-2\\ t_2=-1}} = 1.$$


= OSCILLATORY COVARIOGRAM
EXPLOSIVE OSCILLATIONS
EXPLOSIVE GROWTH
$t_2 = 1 + t_1$
$t_1 + t_2 = 1$
1.0
t_2
TROUGH IN SPECTRUM
$t_1(1-t_2) = -4t_2$
$t_1(1-t_2) = 4t_2$
2.0
1.0
1.0
2.0
t_1
NO TROUGH OR PEAK IN SPECTRUM
PEAK IN SPECTRUM
$t_1^2 + 4t_2^2 = 0$
-1.0
$t_2 = -1$
EXPLOSIVE OSCILLATIONS

2.1a

Such calculations show that the boundaries of region (37) are as depicted in Figure 4. To be in region (37) with $t_2 < 1$ (a requirement of covariance stationarity) implies that the roots of the difference equation are complex. However, complex roots don't imply that (37) is satisfied. Consequently, our two definitions of the business cycle aren't quite equivalent.

The "Slutsky Effect" and Kuznets' Transformation

In the above examples, we have seen that if

$$(40) \qquad y_t = B(L)\varepsilon_t,$$

where ε_t is white noise, then the spectrum of y is related to the spectrum of ε_t by

$$g_y(e^{-iw}) = B(e^{-iw})B(e^{iw})\ \sigma_\varepsilon^2$$

or

$$(41) \qquad g_y(e^{-iw}) = B(e^{-iw})B(e^{iw})g_\varepsilon(e^{-iw})$$

since for the white noise ε, $g_y(e^{-iw}) = \sigma_\varepsilon^2$. It is straightforward to show that for any ε_t, not necessarily a white one, affecting y via (40), the spectrum of y is related to the spectrum of ε by (41). Thus assume that y is related to X by

$$(42) \qquad y_t = \sum_{s=-p}^{q} b_s X_{t-s} \equiv B(L)X_t \qquad\qquad p \geq 0,\ q \geq 0$$

and that the spectrum of X is defined. From (42) we know that

$$y_t y_{t-j} = \sum_{s=-p}^{q} b_s X_{t-s} \sum_{r=-p}^{q} b_r X_{t-j-r}$$

$$= \sum_{s=-p}^{q} \sum_{r=-p}^{q} b_s b_r X_{t-s} X_{t-j-r}.$$

Taking expected values on both sides gives

$$c_y(j) = E(y_t y_{t-j}) = \sum_{s=-p}^{q} \sum_{r=-p}^{q} b_s b_r c_x(j+r-s).$$

The spectrum of y is defined as

$$g_y(e^{-iw}) = \sum_{k=-\infty}^{\infty} c_y(k) e^{-iwk}$$

(43) $$= \sum_{k=-\infty}^{\infty} \sum_{s=-p}^{\infty} \sum_{r=-p}^{\infty} b_r b_s c_x(k+r-s) e^{-iwk}.$$

Define the index $h = k + r - s$, so that $k = h - r + s$. Notice that

(44) $$e^{-iwk} = e^{-iw(h-r+s)} = e^{-iwh} e^{-iws} e^{iwr}.$$

Substituting (44) into (43) gives

$$g_y(e^{-iw}) = \sum_{r=-p}^{q} b_r e^{iwr} \sum_{s=-p}^{q} b_s e^{-iws} \sum_{h=-\infty}^{\infty} c_x(h) e^{-iwh}$$

(45) $$g_y(e^{-iw}) = B(e^{iw}) B(e^{-iw}) g_x(e^{-iw})$$

or

$$g_y(e^{-iw}) = |B(e^{-iw})|^2 g_x(e^{-iw}),$$

which shows that the spectrum of the "output" y equals the spectrum of the "input" x multiplied by the real number $B(e^{iw})B(e^{-iw})$.

Relation (45) can be used to show the famous "Slutsky" effect. Slutsky considered the effects of starting with a white noise ε_t, taking a 2 period moving sum n times, and then taking first differences m times. That is, Slutsky considered forming the series

$$Z_t = (1+L)(1+L)\ldots(1+L)\varepsilon_t = (1+L)^n\varepsilon_t$$

and

$$y_t = (1-L)(1-L)\ldots(1-L)Z_t = (1-L)^m Z_t$$

(46) $$y_t = (1+L)^n(1-L)^m\varepsilon_t .$$

Applying (45) to (46) we have

$$g_y(e^{-iw}) = (1+e^{iw})^n(1+e^{-iw})^n(1-e^{iw})^m(1-e^{-iw})^m\sigma_\varepsilon^2$$

$$= [(1+e^{iw})(1+e^{-iw})]^n[(1-e^{-iw})(1-e^{iw})]^m \quad \sigma_\varepsilon^2$$

$$= [(2+(e^{iw}+e^{-iw})]^n[(2-(e^{iw}+e^{-iw})]^m \quad \sigma_\varepsilon^2$$

(47) $$g_y(e^{-iw}) = \sigma_\varepsilon^2 2^n[1+\cos w]^n \; 2^m[1-\cos w]^m .$$

Consider first the special case where m = n. Then (47) becomes

$$g_y(e^{-iw}) = \sigma_\varepsilon^2 4^n[1- \cos^2 w]^n$$

(48) $$= \sigma_\varepsilon^2 4^n[\sin^2 w]^n .$$

On $[0,\pi]$, the spectrum of y has a peak, at $w = \pi/2$, since there $\sin w = 1$. Notice that since $\sin w \leq 1$, (48) implies that as n becomes large, the peak in the spectrum of y at $\pi/2$ becomes sharp. In the limit, as $n \to \infty$, the spectrum of y becomes a "spike" at $\pi/2$, which means that y behaves like a cosine of angular frequency $\pi/2$.

Similar behavior results for fixed m/n as n becomes large where $m \neq n$. Consider (47) and set $dg_y(e^{-iw})/dw$ equal to zero in order to locate the peak in the spectrum:

$$\frac{dg_y}{dw} = \sigma_\varepsilon^2 2^{m+n}\{ n[1-\cos w]^m [1+\cos w]^{n-1} (-\sin w)$$

$$+ m(1-\cos w)^{m-1} (\sin w)[1+\cos w]^n \}$$

$$= \sigma_\varepsilon^2 2^{m+n} \sin w \{(1-\cos w)^{m-1} (1+\cos w)^{n-1})$$

$$\cdot [m(1+\cos w)-n(1-\cos w)] \}.$$

This expression can equal zero on $(0,\pi)$ only if the expression in brackets equals zero:

$$m(1+\cos w)-n(1-\cos w) = 0$$

which implies

$$\cos w = \frac{1 - \frac{m}{n}}{1 + \frac{m}{n}} \quad ,$$

or

$$w = \cos^{-1}(\frac{1-m/n}{1+m/n})$$

which tells us the frequency at which the spectrum of y attains a peak. For fixed m/n, the spectrum of y approaches a spike as $n \to \infty$. This means that as $n \to \infty$, y tends to behave more and more like a cosine of angular frequency $\cos^{-1}((1-m/n)/(1+m/n))$.

What Slutsky showed, then, is that by successively summing and then successively differencing a serially uncorrelated or "white noise" process ε_t, a series with "cycles" is obtained.

Another use of (45) is in the analysis of transformations that have been applied to data. An example is Howrey's analysis of the transformations used by Kuznets. Data constructed by Kuznets have been inspected to verify the existence of "long swings," long cycles in economic activity of around twenty years. Before analysis, however, Kuznets subjected the data to two transformations. First, he took a five year moving average:

$$Z_t = \frac{1}{5}[L^{-2}+L^{-1}+1+L+L^2]X_t \equiv A(L)X_t.$$

Then he took the centered first difference of the (nonoverlapping) five year moving average:

$$y_t = Z_{t+5} - Z_{t-5} = [L^{-5}-L^5]Z_t = B(L)Z_t .$$

So we have that the y's are related to the X's by

$$y_t = \frac{1}{5}[L^{-5}-L^5][L^{-2}+L^{-1}+1+L+L^2]X_t$$

$$= A(L)B(L)X_t .$$

The spectrum of y is related to the spectrum of X by

$$(49) \qquad g_y(e^{-iw}) = A(e^{-iw})A(e^{iw})B(e^{-iw})B(e^{iw})g_x(e^{-iw}) .$$

We have

$$A(e^{-iw}) = \frac{1}{5}\sum_{j=-2}^{2} e^{-iwj} = \frac{1}{5}\frac{(e^{iw2}-e^{-iw3})}{(1-e^{-iw})} .$$

Thus,

$$A(e^{-iw})A(e^{iw}) = \frac{(\frac{1}{5})^2(e^{iw2}-e^{-iw3})(e^{-iw2}-e^{iw3})}{(1-e^{-iw})(1-e^{iw})}$$

$$= \frac{(\frac{1}{5})^2(2-(e^{iw5}+e^{-iw5}))}{(2-(e^{iw}+e^{-iw}))}$$

$$= \frac{(\frac{1}{5})^2 2(1-\cos 5w)}{2(1-\cos w)} = \frac{(\frac{1}{5})^2(1-\cos 5w)}{(1-\cos w)} .$$

Next, we have

$$B(e^{-iw}) = (e^{+iw5}-e^{-iw5})$$

so that

$$B(e^{-iw})B(e^{iw}) = (e^{iw5}-e^{-iw5})(e^{-iw5}-e^{iw5})$$

$$= (2-(e^{iw10}+e^{-iw10})) = 2(1-\cos 10w).$$

So it follows from (49) that

$$g_y(e^{-iw}) = [\frac{(\frac{1}{5})^2(1-\cos 5w)2}{(1-\cos w)} (1-\cos 10\ w)]\ g_x(e^{-iw})$$

$$= G(w)g_x(e^{-iw}).$$

where $G(w) = 2[(\frac{1}{5})^2(1-\cos 5w)(1-\cos 10w)/(1-\cos w)]$.

The term G(w) is graphed in Figure ___. It has zeroes at values where $\cos 5w = 1$ and where $\cos 10w = 1$. The first condition occurs on $[0,\pi]$ where

$$5w = 0,\ 2\pi,\ 4\pi,$$

or

$$w = 0,\ \tfrac{2}{5}\pi,\ \tfrac{4}{5}\pi.$$

The condition cos 10w = 1 on $[0,\pi]$ where

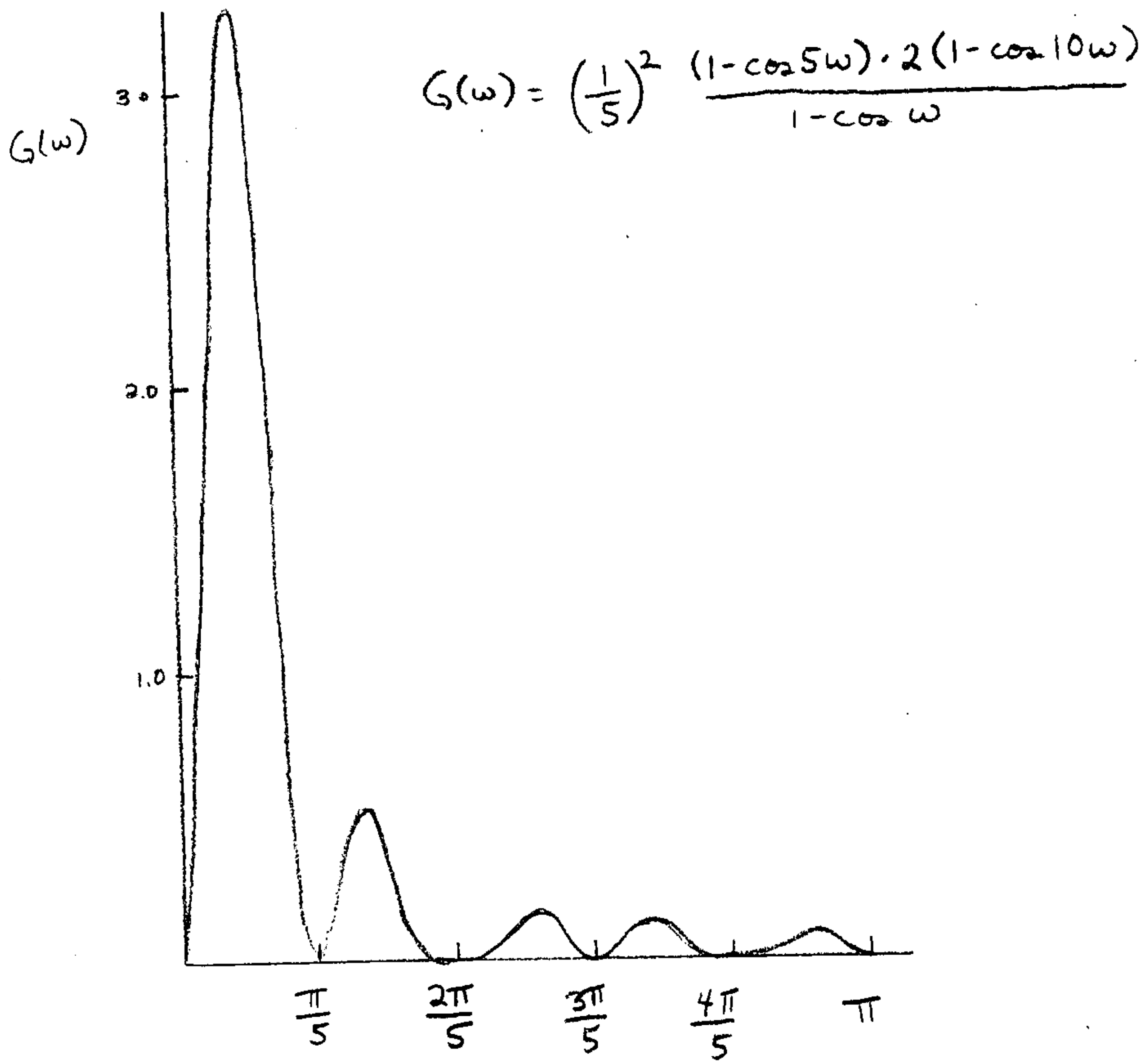
$$10w = 0,\ 2\pi,\ 4\pi,\ 6\pi,\ 8\pi,\ 10\pi$$

or

$$w = 0,\ \frac{1}{5}\pi,\ \frac{2}{5}\pi,\ \frac{4}{5}\pi,\ \text{and } \pi.$$

So G(w) has zeroes at $w = 0,\ \pi/5,\ 2/5\pi,\ 3\pi/5,\ 4\pi/5$, and π.

From the graph of G(w), it follows that even if X_t is a white noise, a y series generated by applying Kuznets' transformations will have a large peak at a low frequency, and hence will seem to be characterized by "long swings." These long swings are clearly a statistical artifact; that is, they are something induced in the data by the transformation applied and not really a characteristic of the economic system. With annual data, the biggest peak in Figure __ corresponds to a cycle of about 20 1/4 years which is close to the 20 year cycle found by Kuznets. Howrey's observations naturally raise questions about the authenticity of the long swings identified by studying the data used by Kuznets.


$G(\omega) = \left(\frac{1}{5}\right)^2 \frac{(1-\cos 5\omega)\cdot 2(1-\cos 10\omega)}{1-\cos\omega}$
G(ω)
3.0
2.0
1.0
π/5
2π/5
3π/5
4π/5
π