Costas Loop

Modules: Sequence Generator, Digital Utilities, VCO, Quadrature Utilities (2), Phase Shifter, Tuneable LPF (2), Multiplier

0 Pre-Laboratory Reading

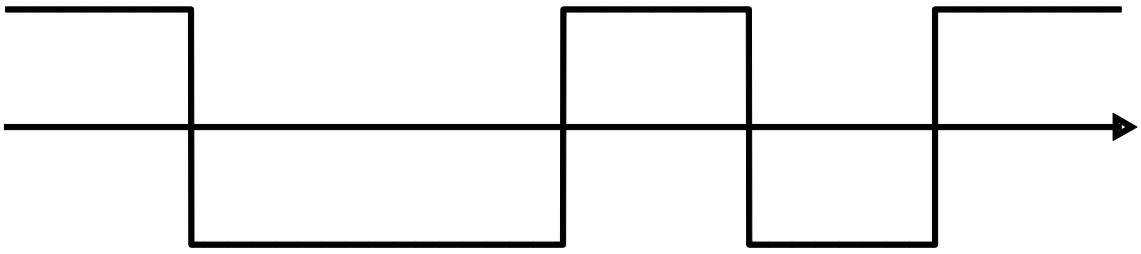
Phase-shift keying that employs two discrete phases (0 and π radians) is often called binary phase-shift keying (BPSK).

0.1 Binary Phase-Shift Keying

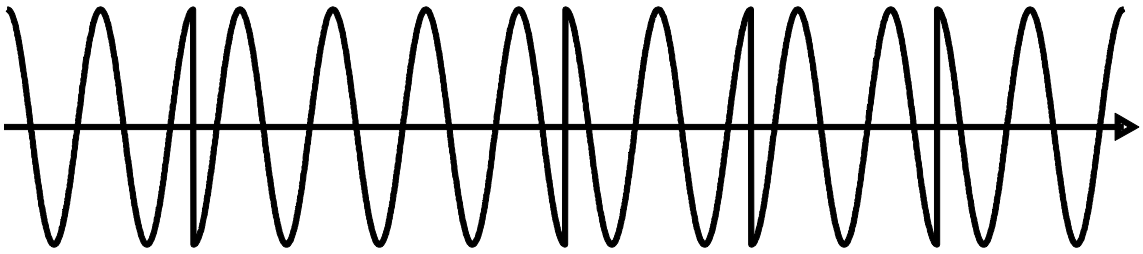
BPSK has the mathematical form

$$y(t) = d(t)\cos(2\pi f_c t) \tag{1}$$

where $d(t) = \pm 1$ is the sequence of bipolar voltages representing the data and f_c is the carrier frequency. During a bit period in which the polarity is $+1$, the carrier has its nominal phase. During a bit period in which the polarity is -1, the carrier phase is different from the nominal by 180°. At those points in time corresponding to a change in bit there is a phase shift of 180°.

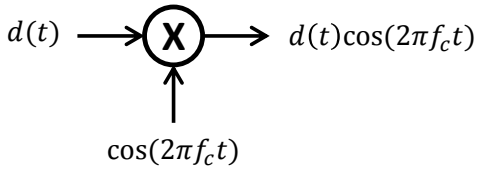


$d(t)$: sequence of bipolar voltages representing the data



BPSK carrier

A BPSK modulator can be implemented (for a relatively small f_c) with a multiplier.



At the receiver the data can be recovered with synchronous demodulation. If a stolen carrier is available, the received signal is multiplied by this stolen carrier. It is assumed here that the stolen carrier is $2\cos(2\pi f_c t)$. If this is the case, then the signal processing in the receiver is

$$\mathcal{S}\{d(t)\cos(2\pi f_c t) \cdot 2\cos(2\pi f_c t)\} = d(t) \tag{2}$$

where $\mathcal{S}\{\cdot\}$ represents the lowpass filtering of the multiplier output. In practice, the bandwidth of $d(t)$ is much smaller than the carrier frequency. The filter passes $d(t)$ while blocking the double-frequency term.

In the field, where the receiver is usually remote from the transmitter, no stolen carrier is available. For synchronous demodulation to work, the receiver must somehow reconstruct a copy of the (unmodulated) carrier from the received signal. It is important to note that this reconstructed copy must match the arriving carrier in phase as well as frequency.

As an example of what doesn't work, consider what happens if synchronous demodulation is attempted with a copy of the (unmodulated) carrier that is offset in phase from the arriving carrier by $\pi/2$ radians. An example of this is $2\sin(2\pi f_c t)$. From trigonometry,

$$d(t)\cos(2\pi f_c t) \cdot 2\sin(2\pi f_c t) = d(t)\sin(4\pi f_c t) \tag{3}$$

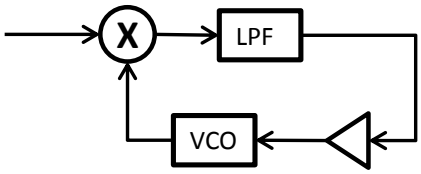
The difference-frequency term is absent in this case; only a double-frequency term is present. Therefore,

$$\mathcal{S}\{d(t)\cos(2\pi f_c t) \cdot 2\sin(2\pi f_c t)\} = 0 \tag{4}$$

In this case, nothing passes the filter and the demodulation fails. This demonstrates the importance of getting the phase, as well as the frequency, right in the receiver. This is known as carrier synchronization. When BPSK is employed, carrier synchronization is done in the receiver with a Costas loop.

0.2 Phase-Locked Loop

A simple phase-locked loop is designed to track a sinusoid. The VCO produces a sinusoid. When the loop is tracking properly, this VCO sinusoid and the input sinusoid have the same frequency. The multiplier produces a difference-frequency term and a sum-frequency term, but only the former passes through the lowpass filter. The output of the filter is an error signal, and it is amplified and then placed at the input to the VCO, completing the loop.

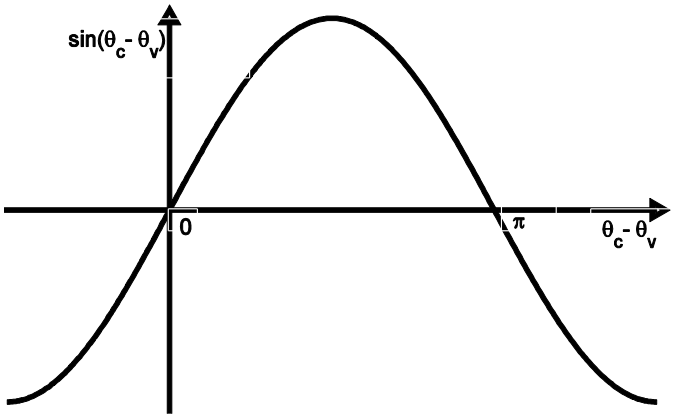


Phase-locked loop

Loop gain is an important parameter in a phase-locked loop. The loop gain is defined as the product of the VCO sensitivity and the amplification in the loop. The behavior of the loop depends on whether the loop gain is positive or negative. In the following discussion, it is assumed that the loop gain for this simple phase-locked loop is positive. (In the TIMS instrument, the VCO sensitivity is negative and the amplifier gain is also negative, so the minus signs cancel and the loop gain is indeed positive.)

With positive loop gain, a positive error signal (appearing at the output of the lowpass filter) causes the VCO output frequency to be larger than its nominal value (the frequency with zero input to the VCO). A negative error signal causes the VCO output frequency to be smaller than its nominal value.

The input to the loop is here modeled as $\sin(2\pi f_c t + \theta_c)$. The VCO output is modeled as $2\cos(2\pi f_c t + \theta_v)$. The multiplier produces a difference-frequency term $\sin(\theta_c - \theta_v)$, and this is the error signal. (The sum-frequency term is blocked by the lowpass filter.) The error signal is plotted below as a function of the phase difference $\theta_c - \theta_v$.



Phase-locked loop: identifying the lock point

There is a stable lock point at the *positive*-going zero crossing: $\theta_c - \theta_v = 0$. The following reasoning shows this to be a point of phase lock. If $\theta_v < \theta_c$, the error signal is positive (assuming $\theta_c - \theta_v$ is not larger than π) and therefore the VCO is forced to produce phase at a faster rate (that is, to produce a larger frequency). This means the feedback action of the loop pushes the loop back toward the point $\theta_c - \theta_v = 0$. If $\theta_v > \theta_c$, the error signal is negative and

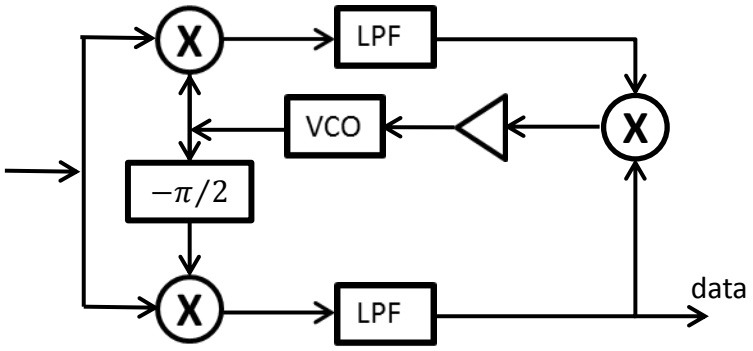
therefore the VCO is forced to produce phase at a slower rate. In this case also, the feedback action pushes the loop back toward the point $\theta_c - \theta_v = 0$. Of course, if $\theta_c - \theta_v = 0$ the error signal is zero, and the loop tends to stay where it is.

It should be noted that a *negative*-going zero crossing is not a stable lock point. If $\theta_c - \theta_v$ moves slightly off the point $\theta_c - \theta_v = \pi$, the feedback action pushes the loop *away* from the point $\theta_c - \theta_v = \pi$. This is called a point of *unstable equilibrium*.

As described above there is one stable lock point per cycle of carrier phase. This point occurs at $\theta_c - \theta_v = 0$. In other words, phase lock corresponds to $\theta_v = \theta_c$. Remember that the input sinusoid is modeled here as a sine and the VCO output is modeled as a cosine. Therefore, when in phase lock (and with positive loop gain) the VCO sinusoid leads the input sinusoid by 90°.

0.3 Costas Loop

A Costas loop is a type of phase-locked loop that is used for carrier synchronization in a receiver when the modulation is BPSK.



Costas loop

Some mathematics will demonstrate how the Costas loop works. The input to the Costas loop is modeled here as

$$\text{Input} = d(t)\sin(2\pi f_c t + \theta_c) \tag{5}$$

where $d(t) = \pm 1$ is the sequence of bipolar voltages representing the data, f_c is the carrier frequency, and θ_c is an implicit function of time representing that part of the total signal phase that is not included in $2\pi f_c t$.

The VCO output is modeled here as

$$\text{VCO Output} = 2\cos(2\pi f_c t + \theta_v) \tag{6}$$

where θ_v is an implicit function of time representing that part of the total signal phase that is not included in $2\pi f_c t$. The output of the filter in the upper channel is

$$K_{\mathrm{UF}} d(t) \sin(\theta_c - \theta_v) \tag{7}$$

where K_{UF} is the DC gain of the upper-channel filter. The term $\sin(\theta_c - \theta_v)$ by itself would be a suitable error signal for a carrier synchronization loop. However, the presence of $d(t)$ means that the signal of Eq. (7) cannot by itself serve as the error signal for the loop.

The local oscillator applied to the lower channel is $2\sin(2\pi f_c t + \theta_v)$. The output of the filter in the lower channel is

$$K_{\mathrm{LF}} d(t) \cos(\theta_c - \theta_v) \tag{8}$$

where K_{LF} is the DC gain of the lower-channel filter.

The third multiplier (the one on the right side of the diagram) produces a suitable error signal:

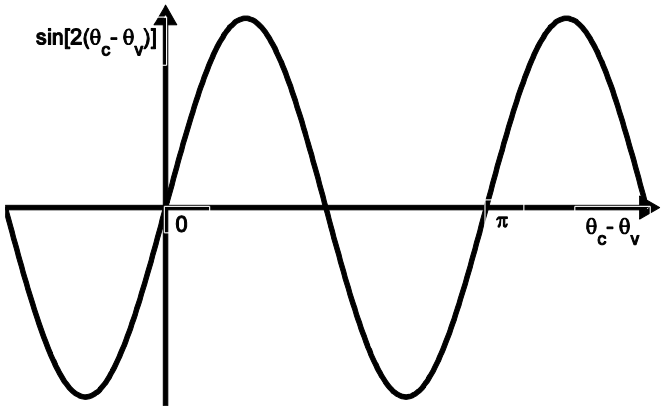
$$\frac{1}{2} K_{\mathrm{UF}} K_{\mathrm{LF}} \sin[2(\theta_c - \theta_v)] \tag{9}$$

Eq. (9) was obtained by noting that $d^2(t) = 1$ and by using the trigonometric identity

$$\sin(\psi)\cos(\psi) = \frac{1}{2}\sin(2\psi) \tag{10}$$

The error signal of Eq. (9) passes through an amplifier with gain G on the way to the VCO input. The loop gain is the product of all gains in the signal path and the VCO sensitivity K_{VCO}. The loop gain is therefore proportional to $GK_{\mathrm{UF}}K_{\mathrm{LF}}K_{\mathrm{VCO}}$. For this experiment G and K_{VCO} are negative and K_{UF} and K_{LF} are positive. Therefore, the loop gain is positive. For a loop with positive loop gain, a positive error signal (the signal at the output of the third multiplier) causes the VCO output frequency to increase and a negative error signal causes the VCO output frequency to decrease.

The error signal for the Costas loop is proportional to $\sin[2(\theta_c - \theta_v)]$. Below that expression is plotted as a function of $\theta_c - \theta_v$.



Costas loop: identifying the lock points

For positive loop gain, the stable lock points are identified as positive-going zero crossings in the above plot. Negative-going zero crossings are *not* stable lock points; instead they are points of unstable equilibrium. For the Costas loop there are two stable lock points per cycle of carrier phase. With positive loop gain, these two stable lock points are at $\theta_c - \theta_v = 0$ and $\theta_c - \theta_v = \pi$. In other words, this loop will lock with either $\theta_v = \theta_c$ or $\theta_v = \theta_c - \pi$. Remember that the input has been modeled as a sine and the VCO output as a cosine. See Eqs. (5) and (6). Therefore, the VCO output will be in phase quadrature with the input.

For example, for a period of time in which $d(t) = +1$ and with the lock point $\theta_v = \theta_c$, the VCO output will lead the modulated carrier by 90°; however, when the bipolar voltage changes so that $d(t) = -1$, the modulated carrier phase changes by 180° so that the VCO then lags by 90°. At the other lock point also, the phase difference between the VCO output and the modulated carrier will be $\pm 90^\circ$, with the sign depending on the polarity of $d(t)$.

When a Costas loop with positive loop gain is in phase lock, the output of the filter in the lower channel will be

$$K_{\mathrm{LF}} d(t) \cos(\theta_c - \theta_v) = K_{\mathrm{LF}} d(t) \tag{11}$$

when the lock point is $\theta_c - \theta_v = 0$ and

$$K_{\mathrm{LF}} d(t) \cos(\theta_c - \theta_v) = -K_{\mathrm{LF}} d(t) \tag{12}$$

when the lock point is $\theta_c - \theta_v = \pi$. In other words, the demodulated data appear either right-side up, $d(t)$, or upside down, $-d(t)$, at the output of the filter in the lower channel. Can we control whether it is $d(t)$ or $-d(t)$? No, we can't. It is essentially random. This is known as a *two-fold phase ambiguity*. This phase ambiguity is not catastrophic, but it does need to be resolved.

In a practical BPSK system, how does the receiver resolve the phase ambiguity? One common method is explained here. In digital communication systems, the data are normally organized into data frames. For example, one data frame might contain the bits for 100 consecutive words, each word representing one sample of the original analog message. Each data frame contains a prefix that contains a short sequence of bits, agreed upon by both transmitter and receiver. This sequence is known as a synchronization word because the receiver can look for it in the arriving stream of bits and thereby identify the start of each new data frame. This process is known as frame synchronization. The receiver can also check to see if the synchronization word (that is expected at the beginning of each data frame) appears right-side up or upside down. If the latter, the receiver knows to invert the entire frame of bits.

What changes if the loop gain is negative? In this experiment the loop gain becomes negative if the amplifier with negative gain is removed from the circuit (since the VCO sensitivity is negative). With a negative loop gain, the stable lock points occur on the plot of $\sin[2(\theta_c - \theta_v)]$

versus $\theta_c - \theta_v$ at negative-going zero crossings. There are two stable lock points per cycle of carrier phase, and these lock points are at $\theta_c - \theta_v = \pi/2$ and $\theta_c - \theta_v = 3\pi/2$. The demodulated data then appear at the output of the filter in the *upper* channel.

When a Costas loop with *negative* loop gain is in phase lock, the output of the filter in the upper channel will be

$$K_{\mathrm{UF}} d(t) \sin(\theta_c - \theta_v) = K_{\mathrm{UF}} d(t) \tag{13}$$

when the lock point is $\theta_c - \theta_v = \pi/2$ and

$$K_{\mathrm{UF}} d(t) \sin(\theta_c - \theta_v) = -K_{\mathrm{UF}} d(t) \tag{14}$$

when the lock point is $\theta_c - \theta_v = 3\pi/2$. As before, the data can appear either right-side up or upside down. This two-fold phase ambiguity can be resolved by the receiver using a synchronization word.

In summary, a Costas loop with negative loop gain behaves in much the same way as a similar Costas loop with positive loop gain. The difference is that the data appear (possibly upside down) at a different location in the circuit.

1 BPSK

The Sequence Generator module will supply the sequence $d(t)$ of bipolar voltages. On the PCB of this module there is a pair of switches in a dual in-line package (DIP). These switches determine the sequences. Set both switches to the up position. You will use the sequence at the analog Y output port.

Connect a $(100/48)$-kHz TTL clock to the clock input of the Sequence Generator. You can create a $(100/48)$-kHz TTL clock from the (100/12)-kHz TTL clock on the Master Signals panel with a divide-by-4 (Digital Utilities).

Place the analog Y output from the Sequence Generator on Channel A. Place the (TTL) sync output of the Sequence Generator on Channel B. Use this same sync output as the trigger source. (A positive trigger level is needed for a TTL trigger source.) The sequence from the Y output is periodic. The sync TTL signal has the same period and therefore permits a stable oscilloscope display. You should observe a stable display of a sequence of bipolar voltages. This will serve as $d(t)$ in this experiment.

Channel A: analog Y output of Sequence Generator
Channel B: TTL sync output of Sequence Generator

Create a BPSK carrier by connecting the analog Y output of the Sequence Generator to one input of the Multiplier and a 100-kHz sinusoid (Master Signals) to the other input. The Sequence

Generator's analog Y output should still be on Channel A. Place the modulator output on Channel B. Use the (TTL) sync output of the Sequence Generator as the trigger source. Consider using the zoom function of the oscilloscope to zoom in on a data transition.

Channel A: analog Y output of Sequence Generator
Channel B: BPSK carrier (modulator output)

Observe the spectrum of the BPSK carrier. You should note that there is no residual carrier.

Channel B: BPSK carrier

Make sure the switch on the VCO module's PCB is set to "VCO". Set the toggle switch on the front panel of the VCO module to "HI". Place the VCO output on the input of the Frequency Counter. With the gain knob on the VCO module set in the fully counterclockwise (zero-gain) position, use the frequency knob to set the VCO output frequency to approximately 100 kHz. Observe the VCO output frequency for a few minutes. It should be clear that the VCO does not have good frequency stability.

Connect the Variable DC output to the VCO input. Adjust the DC source to $+1$ V. Rotate the gain knob on the VCO clockwise until the VCO output frequency decreases by 1 kHz from its nominal frequency. The sensitivity is now set for -1 kHz/V. You should keep the VCO sensitivity at -1 kHz/V for the remainder of this experiment. Disconnect the Variable DC source from the VCO.

The Phase Shifter module has a slide switch on its PCB. Make sure this switch is set to "HI". Adjust the delay of the Phase Shifter so that the phase change is $-\pi/2$ radians for a 100-kHz sinusoid.

Build a Costas loop. Use Quadrature Utilities modules for the three multipliers in this loop. Use the Phase Shifter that you have just calibrated. Use Tuneable LPFs for the upper-channel and lower-channel lowpass filters. Place a Buffer Amplifier between the output of the third multiplier (which multiplies the outputs of the two filters) and the input of the VCO. This amplifier has negative gain and the VCO has negative sensitivity. The minus signs cancel, and the loop gain is positive.

Adjust the bandwidths of the two filters to approximately 26 kHz. (The Tuneable LPF's clock output has a frequency equal to 100 times the bandwidth.) Set the gain knob of each filter to a mid-range position. (The line on this knob should be approximately vertical.)

Connect the BPSK carrier to the Costas loop input. Place a 100-kHz sinusoid (Master Signals) on Channel A. (This should be the same 100-kHz sinusoid that was used in the modulator.) Place a copy of the VCO output on Channel B. Use the Channel A sinusoid as the trigger

source. Connect a copy of the VCO output to the Frequency Counter. Adjust the gain of the Buffer Amplifier until the loop achieves phase lock. If you have any difficulty obtaining phase lock, make slight changes to the VCO nominal frequency (with the loop closed) using the tuning knob on the VCO.

You can recognize phase lock as follows. The Frequency Counter will indicate that the VCO output frequency is a stable 100 kHz. The oscilloscope will have a stable display for the VCO output (as well as for the 100-kHz sinusoid). The stable oscilloscope display is the more reliable indicator of phase lock.


Channel A: 100-kHz sinusoid (the one used within the modulator)
Channel B: VCO output

You should observe that the VCO output is 90° out of phase with the 100-kHz sinusoid from Master Signals. The VCO output may either lead or lag the 100-kHz sinusoid from Master Signals. It depends on which lock point the Costas loop occupies. Take the Costas loop out of lock by changing the Buffer Amplifier gain, then bring the loop back into lock. Repeat this several times. You should find that taking the loop out of lock and then back into lock can (sometimes) cause the lock point to change. You can recognize a change in lock point by noting that the VCO output sometimes *leads* the 100-kHz sinusoid from Master Signals and sometimes *lags* it. (But when the loop is in phase lock, the VCO output and the 100-kHz sinusoid from Master Signals should always be in phase quadrature.)

Place a copy of the BPSK carrier (the input to the Costas loop) on Channel A. Place a copy of the VCO output on Channel B. Use the (TTL) sync output of the Sequence Generator as the external trigger source. You should observe that during some bit intervals the VCO output leads the modulated carrier by 90° and during other bit intervals the VCO output lags by 90°.


Channel A: BPSK carrier
Channel B: VCO output

Keep a copy of the BPSK carrier on Channel A. Place a copy of the local oscillator for the lower channel (that is, the output of the Phase Shifter) on Channel B. Use the (TTL) sync output of the Sequence Generator as the external trigger source.


Channel A: BPSK carrier
Channel B: lower-channel local oscillator (output of Phase Shifter)

You should observe that during some bit intervals the VCO output is in phase with the modulated carrier and during other bit intervals the VCO output is out of phase by 180°. This local oscillator is therefore suitable for a synchronous demodulation of the BPSK carrier.

Place a copy of the analog Y output from the Sequence Generator on Channel A. Place a copy of the lower-channel lowpass filter output on Channel B. Use the (TTL) sync output of the Sequence Generator as the external trigger source.

Channel A: analog Y output of Sequence Generator
Channel B: lower-channel filter output (demodulator output for positive loop gain)

You are now comparing the original data (from the transmitter) to the demodulated data (in the receiver). You should observe that the output of the lower-channel filter is an approximation to either $d(t)$ or $-d(t)$. Take the Costas loop out of lock by changing the gain of the Buffer Amplifier, then bring the loop back into lock. When you do this, the lock point might change. Repeat this several times. You should find that the output of the lower-channel filter is sometimes $d(t)$ and sometimes $-d(t)$, depending on the lock point.

Keep a copy of the Y output of the Sequence Generator on Channel A. Place a copy of the *upper*-channel lowpass filter output on Channel B. Use the (TTL) sync output of the Sequence Generator as the external trigger source. Considering Eq. (7) with $\theta_c - \theta_v = 0$ or $\theta_c - \theta_v = \pi$, you might expect to see zero at the output of the upper-channel filter. However, you will likely see instead a badly attenuated copy of either $d(t)$ or $-d(t)$. The reason for this is that the loop is, at any point in time, *not exactly* on the lock point. If the nominal frequency of the VCO is slightly off 100 kHz, the loop has to work harder to lock and $\theta_c - \theta_v$ is a little off the lock point. Slightly adjust the nominal frequency of the VCO (with the loop closed) and try to minimize the signal at the output of the upper-channel filter. Small adjustments in the gain of the Buffer Amplifier might also help with this minimization.

Channel A: analog Y output of Sequence Generator
Channel B: upper-channel filter output (small for positive loop gain)

So far you have used a Costas loop with positive loop gain. Now you are to experiment with a negative loop gain. Remove the Buffer Amplifier from the loop, so that the loop error signal (from the third multiplier) goes directly to the VCO input. Because the VCO sensitivity is negative and all other gains in the loop are now positive, this new configuration has a negative loop gain.

Adjust the gain of one of the lowpass filters until the loop is in phase lock. Place a copy of the analog Y output of the Sequence Generator on Channel A. Place a copy of the upper-channel lowpass filter output on Channel B. Use the (TTL) sync output of the Sequence Generator as the external trigger source. You should find that either $d(t)$ or $-d(t)$ appears on the output of the upper-channel filter.

Channel A: analog Y output of Sequence Generator
Channel B: upper-channel filter output (demodulator output for negative loop gain)

Take the Costas loop out of lock by changing the gain of one of the lowpass filters, then bring the loop back into lock. Repeat this several times. You should find that taking the loop out of lock and then back into lock can (sometimes) cause the lock point to change. The output of the upper-channel filter will be $d(t)$ for the lock point $\theta_c - \theta_v = \pi/2$ and $-d(t)$ for the lock point $\theta_c - \theta_v = 3\pi/2$.

Keep a copy of the Y output of the Sequence Generator on Channel A. Place a copy of the *lower*-channel lowpass filter output on Channel B. Use the (TTL) sync output of the Sequence Generator as the external trigger source. You will likely see a badly attenuated copy of either $d(t)$ or $-d(t)$ on the output of the lower-channel filter. Slightly adjust the nominal frequency of the VCO (with the loop closed) and try to minimize the signal at the output of the lower-channel filter. Small adjustments in the gain of a filter might also help with this minimization.

Channel A: analog Y output of Sequence Generator
Channel B: lower-channel filter output (small for negative loop gain)

You are now done with the experiment. You should reflect on what you have accomplished in this experiment. A BPSK carrier was demodulated without using a stolen carrier. In practice, with the transmitter and receiver remote from each other, there is no stolen carrier. The receiver must recreate an unmodulated carrier for use in synchronous demodulation. You have shown that a Costas loop accomplishes this. It is all the more remarkable because there is no residual carrier in the transmitted signal. Nonetheless, the frequency and phase information of the (phantom) unmodulated carrier resides within the modulated carrier. The Costas loop extracts this frequency and phase information.