A 56 KILOBAUD RF MODEM

Dale A. Heatherington, WA4DSY

Abstract

This paper describes a 56 kilobaud synchronous RF modem with a 70 kHz bandwidth. The modulation is bandwidth limited MSK generated by a digital state machine driving two digital-to-analog converters, and two double balanced modulators. The carrier phase is shifted plus or minus 90 degrees for each bit. Demodulation is accomplished with a standard quadrature detector chip but various coherent methods can be used for operation at lower signal to noise ratios. The design is relatively simple and easily reproduced.

Design Philosophy and Goals

This modem was designed so that advanced Amateur radio experimenters can duplicate it. Features include:

- No exotic, expensive or hard to find parts.
- Stable, easy to align design.
- Operation on any band allowing 56 kBaud.
- Synchronous design.

Since packet radio uses a synchronous protocol, the modem was designed to be synchronous also. Characteristics of this synchronous modem are as follows:

- Modem supplies transmitter clock.
- Modem receiver recovers and supplies clock and data.
- Modem does NRZ-NRZI conversions.
- Modem scrambles data prior to transmission and descrambles it after reception.

As a result, the synchronous serial interface in the TNC doesn't need a baud rate generator, clock recovery circuit, or a NRZI to NRZ converter.

To simplify the design and construction and to allow operation on various bands, the modem operates at low power (1 mW) in the 28-30 mHz range. The final operating frequency and power output are obtained by selecting an appropriate transverter module. Good performance has been obtained with the Microwave Modules MMT 220/28s and MMT 432/28s. These transverters provide 5 to 7 watts of output power in the 220 or 432 mHz band. These units require a small modification to improve their transmit/receive switching time. A capacitor in the T/R switching circuit must be changed or removed.

Modulation

Many different modulation methods were tried in order to find one with the most desirable characteristics. BPSK was not used primarily because it has large amplitude variations (zero to full power) and uses more bandwidth than the chosen method. Amplitude variations are a problem because they must be passed unchanged by the transmitter. Any nonlinear amplifier stages will cause spreading of the spectrum and interference to adjacent channels. This effect is the same as "flat topping" in SSB transmitters.

FSK was not used because it cannot be demodulated with a coherent demodulator and it is difficult to design and build a stable frequency modulated oscillator capable of large linear frequency shifts (28 kHz).

The chosen method is a bandwidth limited form of minimum shift keying (MSK) . It is slightly different from the ordinary textbook example of minimun shift keying. Minimum shift keying is basically FSK with precise control of several parameters. The frequency shift is exactly 1/4 the baud rate and the phase of the carrier shifts exactly 90 degrees during each baud interval. The amplitude is constant. Unfortunately. MSK produces many unnecessary sidebands which make the signal very wide. When MSK is filtered with a bandpass filter to eliminate the unwanted sidebands, the carrier phase no longer changes by exactly 90 degrees and the amplitude is no longer constant.

The method used in this design removes the unnecessary sidebands and maintains the 90 degree phase shift during each baud interval at the expense of about 3.5 dB of amplitude variation. Also, when the signal is detected using any kind of FM demodulator, the high frequency components appear to be boosted. This necessitates the use of a simple de-emphasis network following the FM demodulator. This network also reduces high frequency noise which results in improved performance.

Bandwidth limited MSK characteristics include:

26 dB bandwidth is 1.25 Hz/Baud.

Uses slightly less bandwidth than BPSK.

Error rate vs carrier to noise ratio performance comparable to BPSK when a coherent demodulator is used.

Has much less amplitude fluctuation than BPSK.

Can be demodulated with several types of detectors.

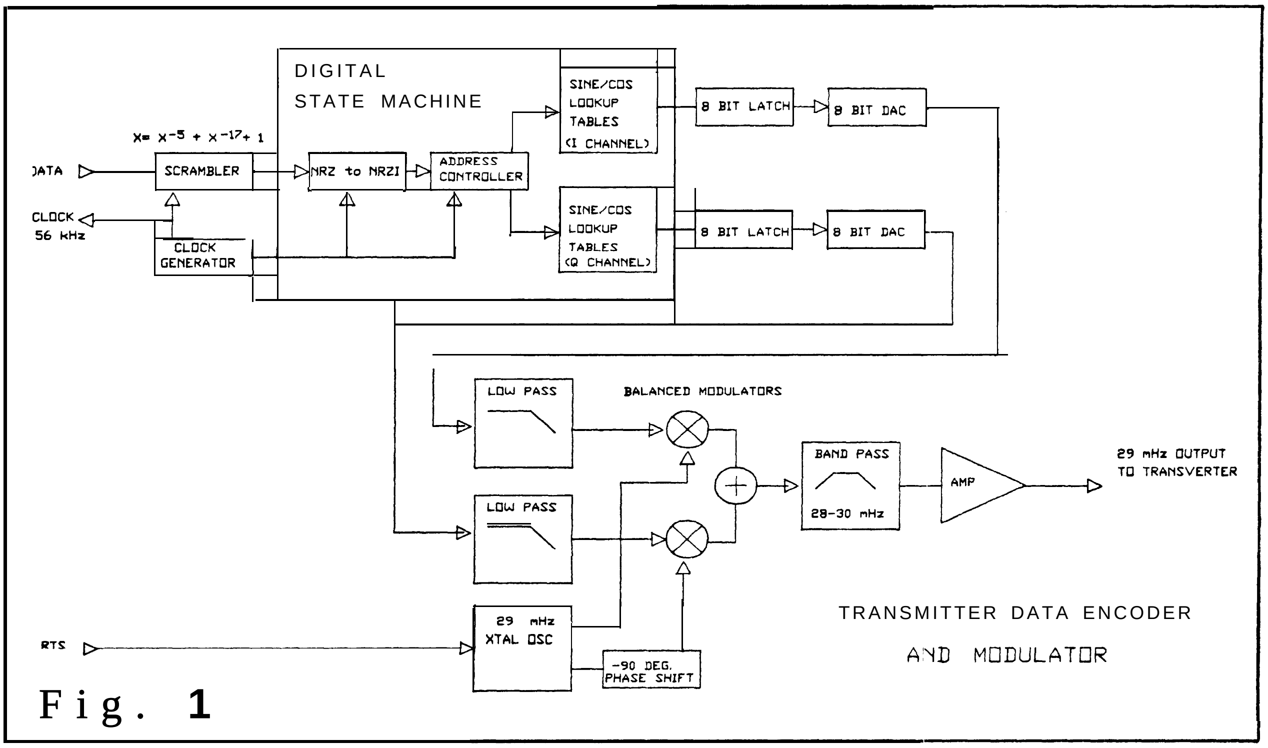
- quadrature detector
- discriminator
- differential phase detector
- Costas loop

Modulation Hardware

Modulation is accomplished with two double balanced modulator chips (type 1496). One modulator is called the "I" (in phase) modulator and the other is call the "Q" (quadrature) modulator. (See Fig. 1) . The carrier frequency is generated **by a** crystal oscillator operating in the 28 to 30 mHz range. The carrier drives the I modulator directly but is phase shifted by 90 degrees before driving the Q modulator. Modulation waveforms are stored in an EPROM chip. These waveforms are read out by a digital state machine into two digital to analog converters (type DAC-08). The analog waveforms are filtered with simple three pole active low pass filters to remove digital sampling noise before being sent to the I and Q modulators. The outputs of the modulators are combined, amplified, and output to -the transverter module. Signals genera-ted this way are unconditionally stable in terms of phase shift and frequency deviation.

Almost any modulation type can be generated with this hardware configuration because the digital state machine has complete control of the amplitude and phase of the carrier. The modulation characteristics are defined by the data stored in the EPROM. The same EPROM also contains the code to run the state machine and NRZ to NRZI converter.

Data rates from 1 to about 120 kilobaud are easily generated by changing the baud rate crystal and/or the divide ratio in the baud rate circuit. Six resistors in the low pass filters also need to be changed for different data rates.



TRANSMITTER DATA ENCODER AND MODULATOR

Fig. 1

Demodulation Hardware

Several methods of demodulation can be used. Since the carrier phase changes **by** exactly 90 degrees during each baud interval, a Costas loop demodulator could be used. The signal also has a frequency shift of 1/4 the baud rate which allows conventional FM or FSK demodulators to be used.

Costas Loop

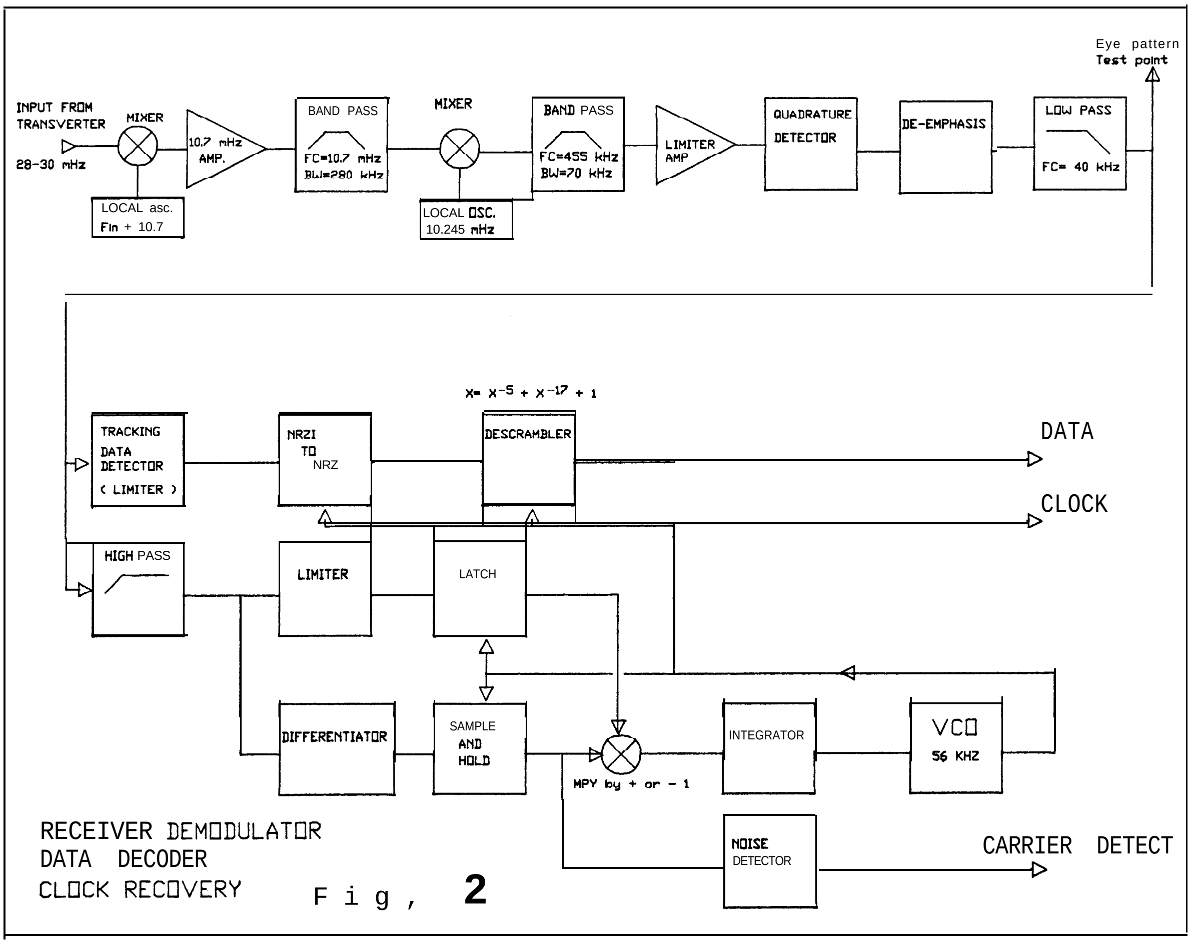
The Costas loop is an intelligent phase lock loop which locks an oscillator to the phase and frequency of the received carrier even with the presence of phase modulation. The received signal is multiplied by the locally generated carrier to recover the original modulation. The main advantage of the Costas loop is its ability to operate at low signal to noise ratios. The disadvantages are its complexity and slow lock-up time. A Costas loop demodulator was built for this project but temporarily shelved because of its 60 ms lock time and large number of parts. It did have a 5 dB S/N advantage over the quadrature detector that will be described below.

Quadrature Detector

In the interest of cost, simplicity, and fast signal acquisition, a conventional quadrature FM demodulator was used in this design (see fig. 2) . The Motorola MC3359 chip was chosen **for** this task. It's more than just an FM demodulator. It also includes an oscillator, mixer, limiter, and several other functions which were not used.

Receiver Bandpass Filter

The receiver bandpass filter was a major stumbling block at the beginning of this project. It had to be 70 to 80 kHz wide with low group delay variations. No "off the shelf" filters could be found at any price. Custom crystal filters had very long lead times and high prices. The solution was simple and cheap: a 3 section L-C bandpass filter operating at 455 kHz. It consists of three 50 uH slug tuned coils and 5 capacitors. The cost is about $4.00. Another filter was required for the 10.7 mHz IF stages. Since its major purpose is image rejection, the response shape was not too critical. A Radio Shack FM broadcast band receiver IF filter had the desired characteristics. It's 280 kHz wide, small, and cheap ($1.00).



RECEIVER DEMODULATOR
DATA DECODER
CLOCK RECOVERY Fig, 2

Data and Clock recovery

After demodulation the signal must **be** processed to recover the clock and data. The data is detected with a tracking data detector, then converted from NRZI to NRZ format. It is then descrambled and sent to the output connector on the modem. Clock is recovered with a sampled-derivative phase locked loop circuit. This circuit aligns the active clock edges with the center of the incoming data bits.

Tracking Data Detector

A data detector is basically a analog comparator. Its threshold is set exactly halfway between the voltage level of a "1" and a "0". It outputs a "1" if the input is higher than the threshold and a "0" if it's lower. There is a problem when the carrier frequency of the incoming signal changes. The voltage levels of the ones and zeros change so the threshold is no longer exactly half-**way** between them. This causes an increase in errors. One common solution, which doesn't work very well, is to AC couple the output of the demodulator to the detector, This is fine if the short and long term average of the number of ones and zeros is equal. This ideal condition cannot be guaranteed even if a scrambler is used. A much better solution is to put some intelligence in the detector so that it averages the voltage level of the ones separately from the average of the zeros then subtracts the two averages to obtain the ideal threshold level. This circuit doesn't care about the ratio of ones to zeros as long as there is a reasonable number of each. A scrambler is used to make sure there is a reasonable number of both ones and zeros. The circuit will compute the correct threshold if the input signal carrier frequency is anywhere within the expected range of the ones and zeros, in this case plus or minus 14 kHz. The maximum frequency offset that can be tolerated is actually limited by the bandwidth of the receiver filter. In this implementation the error rate starts to increase slightly with frequency offsets greater than 5 kHz.

Clock Recovery

For the lowest possible error rate the clock recovery must be fast, accurate, and have low jitter. The state machine circuit used in most TNCs today has a large amount of phase jitter and is not suitable for this application. Its main advantage is low parts count (2 chips). The circuit described here is a sampled-derivative phase locked loop. It gets its phase information from the bit centers, not the zero crossings like most circuits in common use. It does this by taking the derivative (rate of change) of the demodulated data. This converts the data peaks (centers) to zero crossings. This works because the rate of change is zero at the peak of the data bit. This signal is then further processed and compared to the phase of the VCO. The result is an error voltage which is then integrated and applied to the VCO control voltage input. The VCO phase locks to the centers of the incoming bits. Lock time is about 5 milliseconds in this implementation.

Carrier Detector

The carrier detector circuit works by measuring the signal to noise ratio of the signal. This is not an easy task because the data resembles random noise in most aspects. This design solves the problem by first taking the derivative of the demodulated signal, then sampling only the zero crossings of the derivative with the active clock edge. Since the rate of change at the bit centers should be zero, the output of the circuit should be zero. Random noise has random rates of change at the sampling instants which cause the circuit to output a random voltage. This noise voltage is rectified, filtered, and sent to a comparator. If the noise voltage is below a preset threshold, the comparator turns on the carrier detect signal at the digital interface and also turns on a data gate which allows the received data to go out to the interface. An unmodulated signal will also trigger this circuit since there would be no noise at the sampling instants to inhibit the comparator.

NRZI to NWZ Converter

NRZ is a data signaling format in which zeros are represented by a certain voltage level and ones by another. NRZI is a signaling format in which zeros are represented by a change in voltage level while ones are indicated by no change. NRZI coded data is not affected inverting the data voltage levels or the mark/space frequencies in the case of FSK. This modem converts the incoming NRZI data to NRZ data with a simple circuit consisting of a "D" Flip Flop and XOR gate.

Descrambler

There are two good reasons a data scrambler was used in this modem. First, it makes the data stream look like a random stream of ones and zeros regardless of the data being transmitted. This characteristic makes the tracking data detector and clock recovery circuits work better. Second, it makes the RF spectrum look and sound like band limited white noise. In other words, the RF energy is spread evenly over the modems bandwidth and shows no single frequency lines regardless of the data being transmitted. Any potential interference to near by channels is limited to an increase in the noise floor instead of squeaks, squawks, and other obnoxious noises. This type of scrambling is also commonly used in high speed syncronous modems for telephone use.

The hardware to implement the scrambler and descrambler is very simple It consists of a 17 bit shift register and two XOR gates. See Fig. 3. Each transmitted bit is the result of the exclusive ORing of the current data bit with the bits transmitted 5 and 17 bits times before. To descramble the data it is only necessary to exclusive OR the current received bit with the previous 5th and 17th bits. If the data consist of all ones, the scrambler will produce a pseudorandom sequence of bits that will repeat after 131,071 clock pulses or every 2.34 seconds at 56 kilobaud.

Some people may complain that scrambling violates the FCC rule concerning codes and ciphers. It does not. The scrambling algorithm is published here and is available to anyone who wants it. For this reason it does not actually obscure the meaning of the data. If all codes and ciphers were illegal we could not use Morse code, ASCII, or NRZI!

Performance

There are several measures of a modem's performance. One of the most important is its bit error rate. Any well designed modem should not produce any errors if the signal has no noise or distortion in it. The true measure of a modem's quality is its performance under weak signal conditions when the signal to noise ratio is low. These modems were tested to determine their error rate performance at various signal levels and frequency offsets. The results are summarized in the table below.

Signal level		Errors per 1 million bits Freq. offsets in kHz		
uv	dBm	0	+5	-5
.71	-110	3620	43261	11280
.79	-109	2736	8490	3110
.89	-108	843	4694	1510
1.0	-107	129	1680	285
1.1	-106	29	536	77
1.2	-105	0	240	19
1.4	-104	0	23	3
1.5	-103	0	0	0
1.7	-102	0	0	0

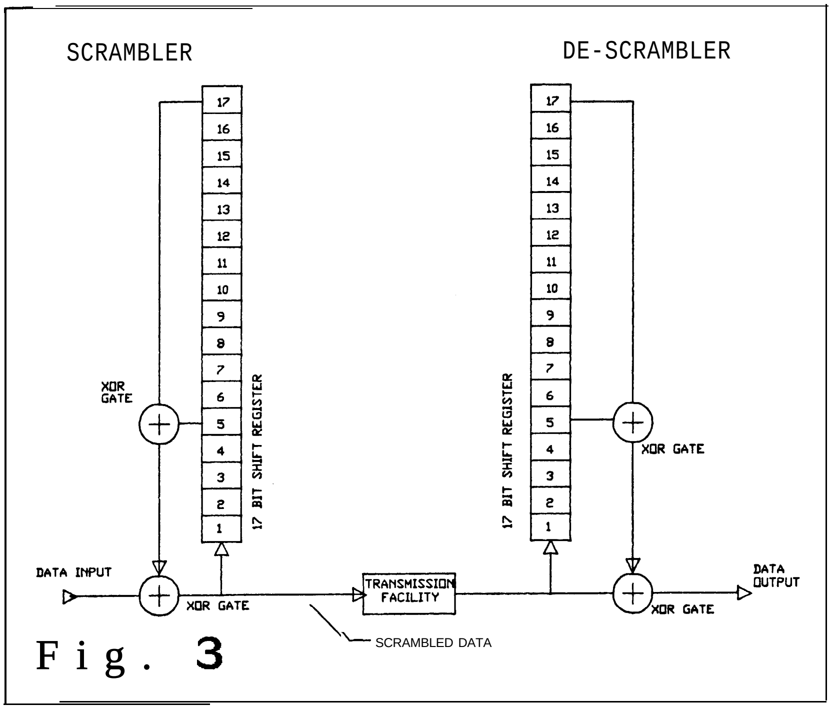


Fig. 3

The table shows that 1.5 microvolts of signal are necessary to achieve an error rate less than 1 per million over a plus or minus 5 kHz frequency range. It also shows that the performance is degraded by 2 dB if the frequency is offset by 5 kHz.

This data was obtained using a MMT 432/28 transverter without a GASFET preamp. The published noise figure for this unit is 3 dB. Adding a pre-amp should improve the performance.

Another important performance parameter is the delay time from transmitter turn-on to valid data at the receiver. This modem requires a 10 to 13 millisecond delay after the transmitter is keyed before data can be transmitted. 5 to 7 ms of this time is required for the crystal oscillator to start and stabilize. The remaining 5 to 6 ms allows the distant receiving modem time to phase lock its clock to the incoming signal and detect the carrier. This time delay is not as low as it should be and work is being done to reduce it. If the delay were zero a 256 byte packet should take 36.5 ms to transmit. This modem takes up to 50 ms of transmission time or 36% overhead. This overhead percentage could be reduced by transmitting longer packets.

Applications

So, what do you do with a 56,000 baud RF modem? One obvious application which has received much press lately is network backbones. One popular opinion seems to be that high speed modems will solve network congestion problems. This assumes that current firmware and software will run at 56 kilobaud. It doesn't! One of the major problems of this project was getting packet software to operate at this speed. No TNC running AX.25 will operate at 56 kilobaud. IBM PCS running at 8 mHz can only handle serial port data at 38.4 kilobaud or less without dropping characters. To conduct on the air tests it was necessary make major modifications to a Z-80 assembly language program called KISS-TNC written by K3MC. This program resides in an EPROM in the TNC. Its main job is to convert SYNC frames to ASYNC frames and send them to the serial port on a PC. A program running in the PC is responsible for doing the protocol, in this case NET.EXE by KA9Q, which implements TCP/IP. Although KISS-TNC was successfully modified for 56 kilobaud, NET.EXE would not run any faster than 19.2 kilobaud. The result was that the PC to TNC interface ran at 19,200 baud while the TNC to RF modem interface ran at 56,000 baud. Needless to say, NET.EXE could not keep the channel busy. This may not be a problem because the channel is a shared resource and should not be hogged by one user anyway.

Digital Voice

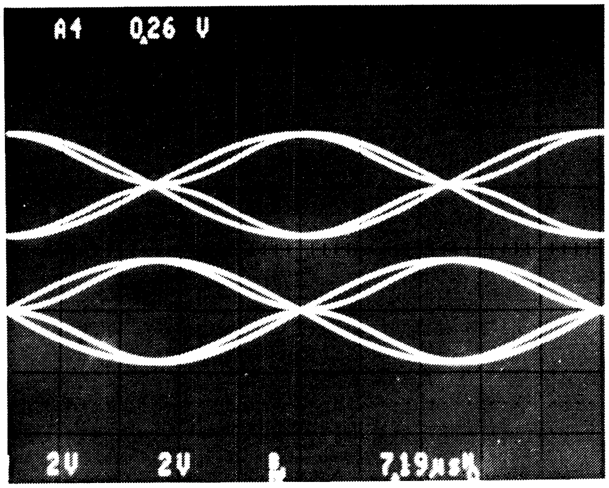
This modem is fast enough to carry real time digitized speech. Two methods of converting speech to digital format are popular. The phone company uses Pulse Code Modulation (PCM) at a 64 kilobaud data rate. Unfortunatly someone set the upper limit for amateur digital communication at only 56 kilobaud thus preventing the use of many cheap CODEC chips in the market today. They only work at 64 KBPS. One good alterative to PCM is continuously variable slope delta modulation (CVSD). Distorted but intelligible speech can be transmitted at speeds as low as 9600 BPS with CVSD. At 56 KBPS it sounds as good as typical communications quality FM. The CVSD chip used in the experiments with this modem is the Motorola MC34115. Its price is in the $2.00 range. CVSD also works well in the presence of data errors. Digitized speech can be understood with more than 10% bit errors. It's quite noisy at that error rate and sounds like a weak FM signal.

Digital Video

Video frame grabbers are available which digitize and store pictures from a NTSC video source in computer memory. Images stored in this manner can be transmitted digitally. A video frame consisting of 256 by 256 pixels with 64 shades of gray can be transmitted in 7 seconds at 56 kilobaud. The same image would take more than 5 minutes to send with the current 1200 baud standard.

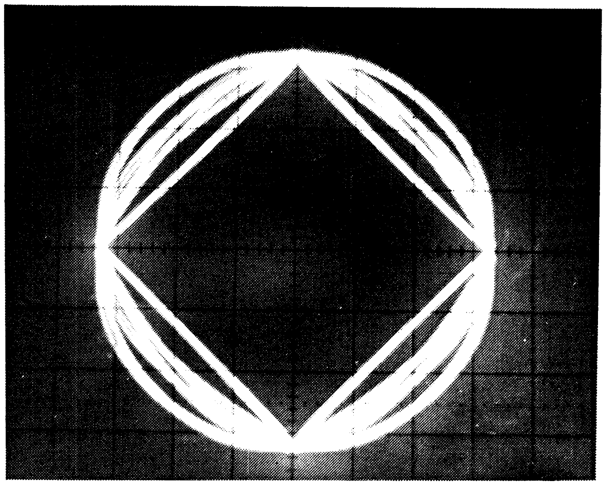
Conclusion

Bandwidth limited MSK is a power efficient modulation method with reasonable bandwidth requirements and low amplitude fluctuations. The hardware required for generation and demodulation is simple and reliable. A wide variety of demodulators can be used depending on the cost/performance tradeoffs. It is an excellent choice for high speed RF data links.



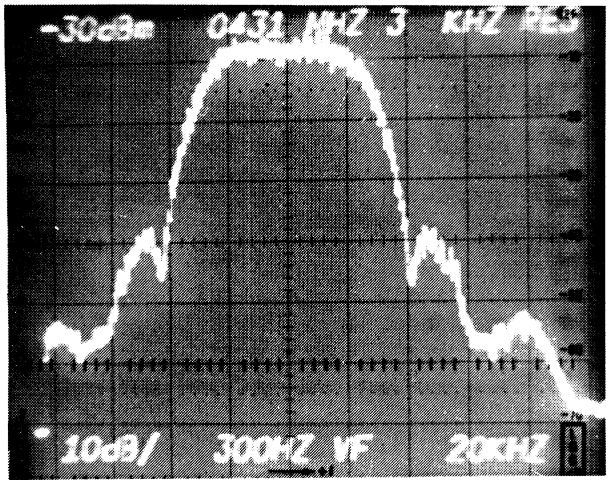
Modulation waveforms

Top trace : I modulation
lower trace : Q modulation

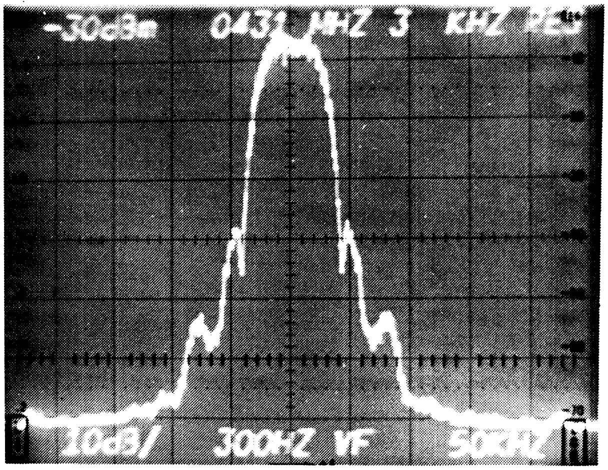


Signal constellation

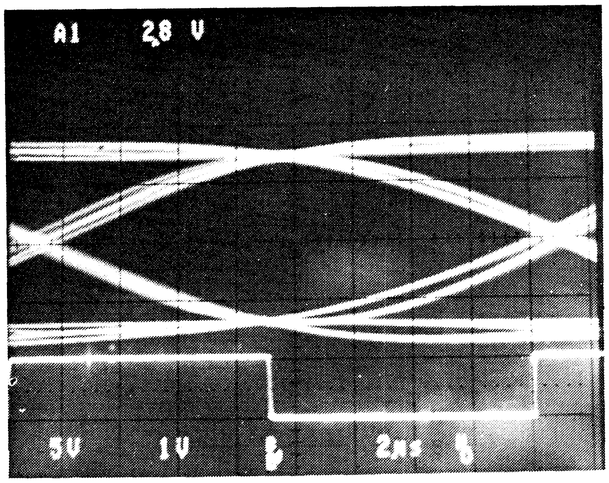
Transmitter X,Y display
X axis : I modulation
Y axis : Q modulation



Bandwidth limited MSK spectrum
20 kHz/div. horz.
10 dB/div. vert.

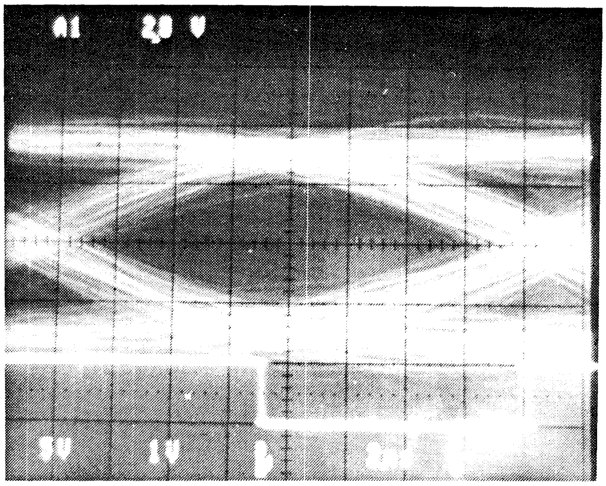


Bandwidth limited MSK spectrum
50 kHz/div. horz.
10 dB/div. vert.



56 kilobaud EYE

Top trace : Receiver EYE pattern
Lower trace : recovered clock



56 kilobaud EYE

Top trace : weak signal EYE pattern
RF signal level is about 1.4 uV
lower trace : recovered clock