AN ABSTRACT OF THE THESIS OF

Jeff D. McNeal for the degree of Masters of Science in

Electrical and Computer Engineering presented on February 11, 1998. Title:

A Comparison of Two Types of Zero-Crossing FM Demodulators for Wireless Receivers.

Abstract approved: ______________________________

Sayfe Kiaei

A comparison of two novel demodulators. The first is a basic zero crossing demodulator, as introduced by Beards. The second is an approach proposed by Hovin. The two demodulators are compared to each other and to the conventional method of demodulation.

A Comparison of Two Types of Zero-Crossing FM Demodulators for Wireless Receivers

by

Jeff D. McNeal

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Masters of Science

Completed February 11, 1998
Commencement June 1998

Masters of Science thesis of Jeff D. McNeal presented on February 11, 1998

APPROVED:

Redacted for Privacy

Co-Major Professor, representing Electrical and Computer Engineering

Redacted for Privacy

Co-Major Professor, representing Electrical and Computer Engineering

Redacted for Privacy

Chair of the Department of Electrical & Computer Engineering

Redacted for Privacy

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Jeff D. McNeal, Author

ACKNOWLEDGMENT

For all of the guidance, help and support, special thanks to all of my advisors and mentors, Sayfe Kiaei, John and Ginny Stonick.

I owe great deal to Roger Traylor, for the help he gave me with the digital design and VHDL.

Most of all thanks to Kathleen, for all the moral support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

To Mark

A COMPARISON OF TWO TYPES OF ZERO-CROSSING FM DEMODULATORS FOR WIRELESS RECEIVERS

1. INTRODUCTION

The purpose of this thesis is to evaluate a new type of zero crossing demodulator in terms of output signal quality, power consumption, and gate count. The new demodulator will be compared to more common forms of FM demodulators, including the traditional zero crossing demodulator. The motivation for the paper is to find an FM demodulator that will work well in a wireless receiver. In order to perform well in a wireless environment, a circuit must not consume large amounts of power, and should have a low gate count, to preserve physical space.

In the first chapter, basic modulation and demodulation will be presented and common forms of FM demodulation will be outlined. Then both the traditional zero crossing demodulator and the new Hovin demodulator will be presented in more detail. The third chapter will analyze both of these zero crossing demodulators and compare them to each other and to a few common types of FM demodulators, including the Phase Locked Loop (PLL) and the ideal demodulator. Finally, conclusions will be presented.

1.1. Modulation

There are two basic ways that information can be encoded in a sinusoidal wave. The first way is to vary the amplitude of the sinusoidal wave as a function of the information-bearing signal. This process is called amplitude modulation or AM. The second basic way is to vary the phase angle of the sinusoid as a function of the information bearing signal. This process is called angle modulation. There are two subtypes of angle modulation, frequency modulation and phase modulation, denoted FM and PM respectively. Phase modulation involves varying the phase angle as a function of the amplitude of the data signal. Frequency modulation involves varying the instantaneous frequency as a function of the amplitude of the data signal. In this thesis we focus on the more commonly employed FM, although with some minor modification FM and PM are interchangeable.

Each of these two classes of modulation, AM and FM, has its own advantages and disadvantages. Amplitude modulation is easier to understand and visualize, and as a result was utilized first (c1910). An AM signal consists of a sinusoid who's amplitude is varied with time as a function of the data signal

$$AM(t) = kx(t)sin(\omega_c t) \tag{1.1}$$

It is fairly simple to modulate and demodulate AM signals, the circuits required are simple and inexpensive to build. They use a fixed bandwidth that can be easily controlled. However the power of an AM signal fluctuates over time, and since they require linear amplification over this range, they are relatively power inefficient to transmit. When the modulating signal has a high amplitude, the power of the modulated wave will be high, when the modulating signal has a low amplitude, the power consumption will be lower.

However even if the modulating signal has no amplitude, the carrier will still have some minimal amplitude, so that the receiver does not loose the signal. This power that is used when there is no input signal contains no information, and so it is wasted.

The information in an AM signal is contained in its envelope. To extract the information correctly, the envelope must not be distorted. This requires, a linear amplifier. The linear high power amplifier must always operate in it's linear range. The result is that most of the time it will not be operating at maximum efficiency, which is a significant waste of power. AM transmitters require the use of a Class A amplifier, who's maximum efficiency is less than 50 percent, while an FM system can use a class C amplifier, who's efficiency can reach nearly 100 percent. [1]

FM signals are more complex to understand and to visualize. FM and PM, popularized by Edwin H. Armstrong, who demonstrated their usefulness in 1933, are defined in equations 1.2 and 1.3.

$$FM(t) = A_c sin\left[\omega_c t + k_f \int_{-\infty}^{t} x(\alpha)\delta\alpha\right] \tag{1.2}$$

$$PM(t) = Asin(\omega_c t + k_p x(t)) \tag{1.3}$$

There are several parameters that are useful when FM systems are being discussed. One is the frequency deviation, or Δf, defined in equation 1.4. Another is the modulation index, which is denoted by β, and is defined in equation 1.6.

$$\Delta f = k_f \times max\,|x(t)| \tag{1.4}$$

For this thesis,

$$\int_{-\infty}^{\alpha} x(t) = sin(\omega_m t) \tag{1.5}$$

and therefore, the maximum value of x(t) will be one. So equation 1.4 simplifies to $\Delta f = k_f$.

$$\beta = \frac{\Delta f}{f_m} \tag{1.6}$$

Where f_m is the maximum frequency present in the message signal.

The circuits for FM transmitters and receivers are more complex to design and build than AM circuits, since the process is not a simple mixing of the message with the carrier. Also, an FM signal has a transmission bandwidth that does not vary linearly as a function of the data signal.

However, there are three major reasons why FM transmission is used instead of AM. The first, and probably most important reason, is that FM signals have a fixed amplitude. The result of this characteristic is that the transmission power is also fixed. Since the output power is lower than an AM system, they are more economical to operate. Second, a FM system is more resistant to noise and interference than an AM system. Since a FM receiver only needs to recover the frequency of the incoming signal, noise that contributes to amplitude variations is not detrimental. Third, in a FM system, the designer has a trade-off between channel bandwidth and signal to noise ratio (SNR) performance. For example if the designer knows that the system will be operating in a noisy environment, he can choose to use more bandwidth, and get better noise performance. In order to increase SNR performance in an AM system, power must be increased.

1.2. Demodulation

There are many different ways to demodulate an FM signal. Each method has different strengths and weaknesses, which make them suited to different applications. For the purposes of this paper, a demodulator that can be easily realized on an IC and does not require a lot of power is desired.

Today some of the most common demodulators are frequency discriminators. A frequency discriminator consists of a slope circuit followed by an envelope detector [2]. This circuit has been used successfully for many years, but is not easy to build in IC form since it requires a transformer and several capacitors.

Another demodulator that has been used to demodulate FM signals is the Phase Locked Loop, or PLL. The PLL is a device that consists of three major components, a multiplier, a loop filter, and a voltage controlled oscillator (VCO), arranged in a negative feedback loop. The mixer and loop filter combine to produce an error signal, which is fed to the VCO, which attempts to match the exact phase of the incoming signal. This error signal is then the message signal output. [2]

The PLL can be realized on an IC, but it is an analog device, and requires a lot of fine tuning to work properly.

A third way to demodulate FM signals is using a zero crossing demodulator. The zero crossing (ZC) demodulator works on the principle that the information in an FM signal, contained in the local frequency of the signal, is equal to the inverse of the period of the signal. Therefore if the period of the signal is measured, that is all the information needed to determine its local frequency. The simplest way to measure the period of a periodic signal is to measure the time between one positive going zero crossing to the next. To do this all that needs to be done is detect the positive going zero crossing, and measure the time elapsed until the next positive going zero crossing is detected. To detect a zero crossing the input signal is compared with a delayed version of itself. If the two signals have the same sign, no zero crossing has occurred. If the signs are different, a zero crossing has occurred. If the first sign is negative and the second sign is positive, then a positive going zero crossing has occurred. Now that a method is known for detecting positive going zero crossings, a method is needed for measuring the elapsed time between them.

The simplest way to determine the time between zero crossings is with a clock that operates at several times frequency of the periodic signal input. When a zero crossing is detected, the counter value is recorded, and the counter is reset so it will be ready measure the number of clocks until the next crossing. When this process is complete, the continuous time waveform has been demodulated into a series of integers representing its period. To convert these period numbers into frequency numbers, they are inverted. Plotting these integers versus time will give an approximation of the data signal that was encoded in the FM waveform.

Since the zero crossing detector is completely digital, it lends itself nicely to being realized on an IC. In this thesis, two methods for measuring the time interval between zero crossings will be examined.

2. ZERO CROSSING DEMODULATORS

Zero crossing demodulators are a good choice for IC implementation, since they are almost completely digital. An all digital design is much more immune to the process variations that can plague analog designers. A zero crossing demodulator does not have the same problems with linearity and distortion that some analog methods of demodulation have. In this chapter I explain in more detail how the traditional zero crossing and Hovin demodulators work. Both of these types of zero crossing demodulators are presented in detail.

2.1. Traditional Zero Crossing Demodulator

The traditional zero crossing demodulator consists of three main parts, which are shown in figure 2.1. The first part is a zero crossing detector, which outputs a pulse at every zero crossing of the incoming signal. The second part is a counter that counts up or down until it receives a pulse from the zero crossing detector, at which time it outputs the current count value and resets. The third part consists of a set of filters to smooth the counter output and produce the demodulated signal.

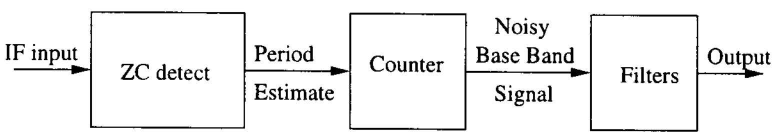

FIGURE 2.1: Zero Crossing Demodulator Parts

2.1.1. *Zero Crossing Detector*

The input to the zero crossing detector is an amplitude limited intermediate frequency (IF) signal. The FM input signal has already been received, band-pass filtered and mixed down to the IF, and is passed through a limiter to produce a $-1 + 1$ binary square wave.

The zero crossing detector produces a stream of pulses, one pulse for each zero crossing. Since information in a FM signal is contained in the instantaneous frequency, as discussed in the last chapter, it can be extracted from the time between zero crossings, ie, the instantaneous period. To generate pulses at each zero crossing, the detector compares the limited IF signal with a delayed version of itself, by passing the IF signal through a flip-flop, which delays it one clock cycle, as shown in Figure 2.2.

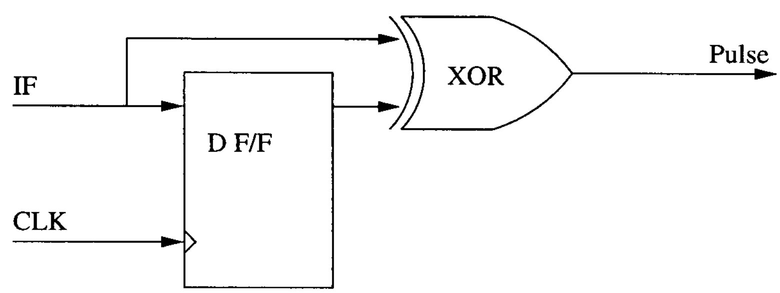

FIGURE 2.2: Zero crossing detector

The detector uses an exclusive-or gate to compare the sign of the signal with a delayed version of itself. When the inputs to the exclusive-or gate are the same its output is zero. When they are different, its output is one. The only time the two versions of the signal will be different is when the signal has crossed through zero between the two samples. Thus the pulse train output of the exclusive-or gate is an approximation of the location of the zero crossings. The amount of delay of the flip-flop determines the minimum pulse width that the exclusive-or gate can produce. Since we want the detector to be as accurate as possible, we want the pulses to be as narrow as possible, ideally delta functions. The way to do this is to increase the clock frequency.

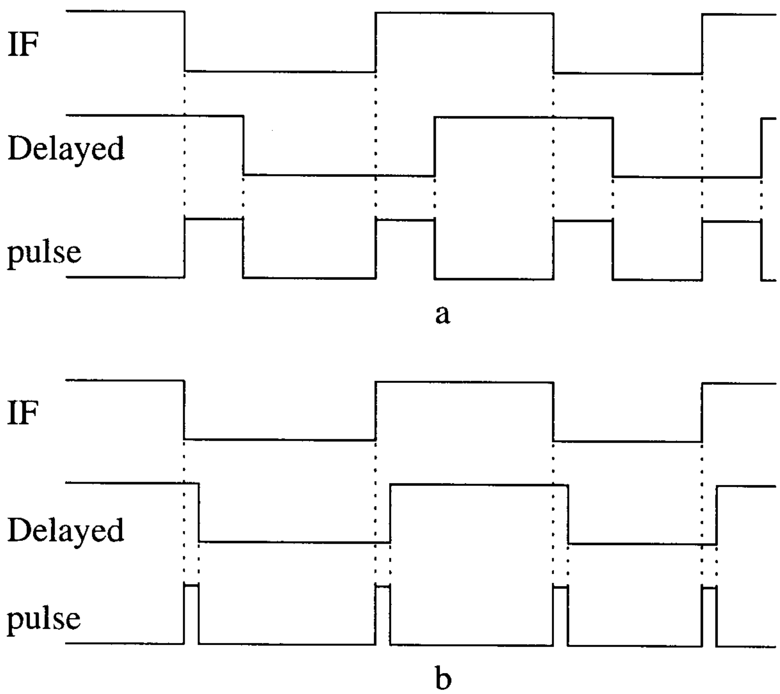

FIGURE 2.3: a) low oversampling rate b) High oversampling rate

The clock frequency of the flip-flop must be more than twice the sum of the IF frequency and bandwidth of the signal for the detector to work, this serves as a lower limit on clock speed. The only upper limit is the maximum clock rate of the flip-flop. To double the effective resolution, both the positive going zero crossings and the negative going zero crossings are counted, i.e. two counts per period. In this paper, all of the demodulators presented and simulated count both positive going zero crossings and negative going zero crossings.

2.1.2. *Counter*

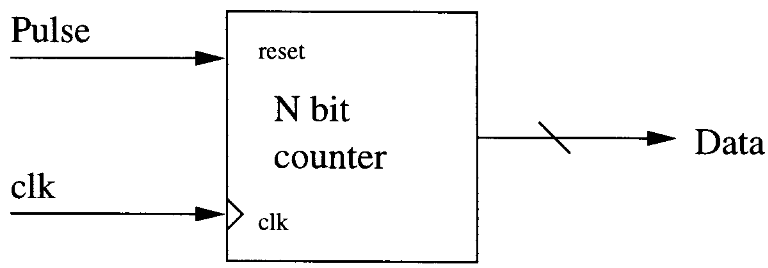

FIGURE 2.4: N bit counter

The output of the counter, figure 2.4, is the number of clock ticks between successive zero crossings. For this counter output to be useful in demodulating the signal, we must be able to use it to differentiate between small changes in frequency. The faster the counter operates, the greater its ability to measure small changes in frequency, leading to higher resolution, Figure 2.3.

To have the highest resolution possible, the counter should be clocked as fast as possible. However, clocking the counter at a high rate requires that it be able to represent very large numbers, requiring it to have more bits. This is an important design trade-off. A high resolution system is desired, but it also must be simple and efficient, since we are focusing on a demodulator for an IC implementation. Later this trade off is examined in more detail. In a CMOS system, power consumption is linearly related to clock speed, since charge is only flowing when a gate changes states. In a synchronous system, such as this one, gates only change states on clock edges. At higher clock rates, charge is flowing more often, so more power is consumed. Since this system is intended for use in a wireless receiver, which is battery operated, power consumption needs to be kept as low as possible.

2.1.3. Filtering

The last element in the demodulator is the filtering section. It is required to perform two functions: First, it removes out of band noise created by the previous sections, acting as a low pass filter. As will be shown later, most of the noise added by the zero crossing detector and counter is higher frequency noise which can be removed using a low pass filter, like a comb filter. Second, it decimates the data stream down to a rate that is consistent with twice the bandwidth of the message signal. The counter is producing data at roughly 100 kHz, and the base band signal is at 16 kHz.

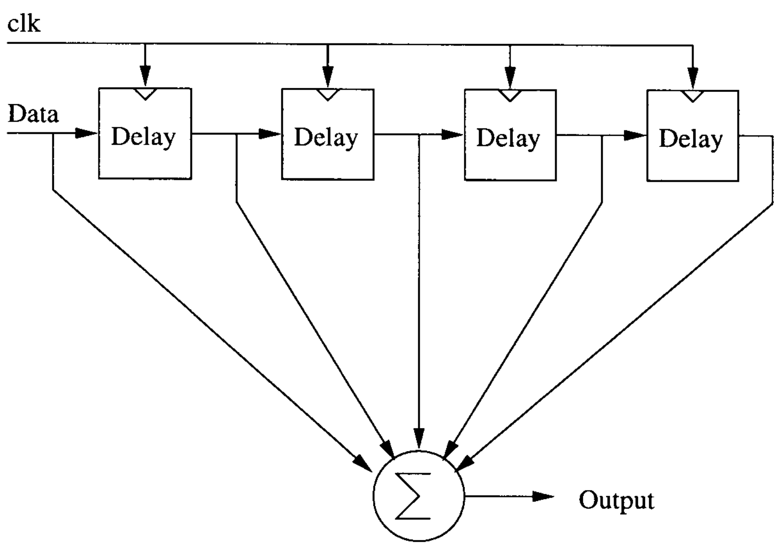

FIGURE 2.5: Example of a comb filter used to decimate and filter

A comb filter achieves both of these goals, (Figure 2.5). It is a low pass filter since it adds many consecutive samples together. It also is a decimator because it only outputs one sample for several inputs. However, a second filter often follows the comb filter, to further reduce the high frequency noise.

One additional problem that needs to be addressed is that the counters output occurs at zero crossings of the IF signal. This causes data to become available at a rate that is not constant, but is a function of the data itself. Therefore as the local frequency of the IF signal changes, the data rate also changes. For example, if the value of count is 12, then the data will be available every 12 clock cycles, if the value of count is 500, the data will only be available every 500 clock cycles. In a synchronous digital system, this can cause several problems. Variable rate data must be synchronized before any meaningful processing can be performed on it. If the data samples are merely delayed to make the rate constant, then their value is no longer correct. This will cause distortion in the output signal, which is demonstrated later. To avoid distortion the data must be re-sampled using a synchronous clock. One way to do this is to use a comb filter structure. The data will be shifted through the delay line as it is produced. Then when a clock edge occurs several samples will be averaged together producing a synchronous, constant rate output.

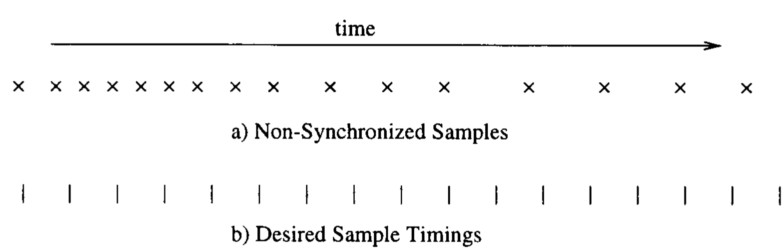

FIGURE 2.6: a) timing of data at the output of the counter b) timing needed for successive circuits

2.1.4. *One Bit Zero Crossing Demodulator*

One bit ZCD is a special case of the basic zero crossing demodulator that makes it even more simple. In the one bit case, the counter section of the demodulator is eliminated, and the pulse stream is used as a virtual count vector. Since the output of the exclusive or gate is 1 if one zero crossing occurred and 0 if no zero crossings occurred, this is an accurate count of zero crossings. This pulse stream is fed into the same set of filters that the regular ZC uses. Since the pulse stream contains the locations of all of the zero crossings, it works the same way that the regular ZC works. Since the one bit method does not need a counter, many gates can be saved, meaning less power consumed. There is a disadvantage to using a one bit zero crossing detector, most of its circuits are going to run at a higher clock frequency. Since the counter in a regular ZC performs some decimation, subsequent circuits receive data at a lower rate, saving power. On the other hand, the one bit demodulator doesn't require any interpolation, since data is coming at a regular rate.

To avoid both the interpolation problem associated with traditional zero crossing demodulators and the high clock rate associated with one bit demodulators, the newer Hovin method can be used.

2.2. Hovin Zero Crossing Demodulator

The new type of zero crossing demodulator being evaluated is the Hovin demodulator [3] [4] [5]. Hovin begins his explanation of the demodulator with the circuit for a sigma delta modulator. The sigma delta modulator takes an analog input signal, and produces a quantized output signal. One of the intermediate steps in the sigma delta modulator involves a step that is very similar to frequency modulation, and Hovin uses this to derive a form of zero crossing demodulator.

The input to the sigma delta modulator, x(n), is a sampled analog signal. The output from the sigma delta, y(n), is a quantized version of the input, x(n).

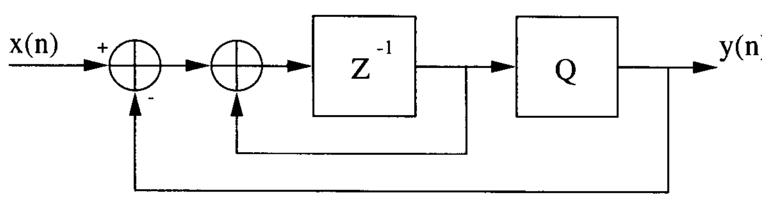

FIGURE 2.7: First order Delta-sigma modulator

The output of the sigma delta circuit in figure 2.7 may be expressed as

$$Y(z) = Q\left[\frac{z^{-1}}{1-z^{-1}}X(z) - \frac{z^{-1}}{1-z^{-1}}Y(z)\right]. \tag{2.1}$$

Given equation 2.2, which is proved in the appendix of [5], the already quantized values Y(z) may be resolved from the quantization function in equation 2.1, resulting in equation 2.3.

$$\sum_i Q[x_i] = Q\left[\sum_i Q[x_i]\right] \tag{2.2}$$

$$Y(z) = Q\left[\frac{z^{-1}}{1-z^{-1}}X(z)\right] - \frac{z^{-1}}{1-z^{-1}}Y(z) \tag{2.3}$$

Collecting Y(z) terms and simplifying results in equation 2.4. This is mathematically equivalent to equation 2.1 but is represented by a different circuit, which is shown in figure 2.8.

$$Y(z) = (1 - z^{-1})Q\left[\frac{z^{-1}}{1 - z^{-1}}X(z)\right] \tag{2.4}$$

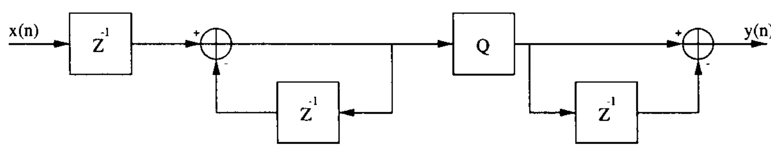

FIGURE 2.8: Hovin integrator differentiator modulator demodulator

Recall from equation 1.2, reprinted here for convenience, that the output of a frequency modulator is equation 2.5. To extract the message signal we need a phase detection and differentiation. The output of a phase detector, figure 2.9 and equation 2.6, is the difference in the phase of its two inputs.

$$FM(t) = Asin(\omega_c t + k_f \int_{-\infty}^{t} x(t)\delta t) \tag{2.5}$$

$$\int x(t)\delta t = P.D.\left[sin(\omega_c t), sin(\omega_c t + \int x(t)\delta t)\right] \tag{2.6}$$

If a message signal is passed through a frequency modulator and then a phase detector, equations 2.5 and 2.6 combine to yield a simple integration of the message signal. Therefore the integration block in figure 2.8 can be replaced with an FM modulator followed by a phase detector. The new circuit is shown in figure 2.10.

This new circuit, an FM modulator followed by a phase detector, a quantizer and finally a differentiator, forms the basis for Hovin's modulator / demodulator. Recalling

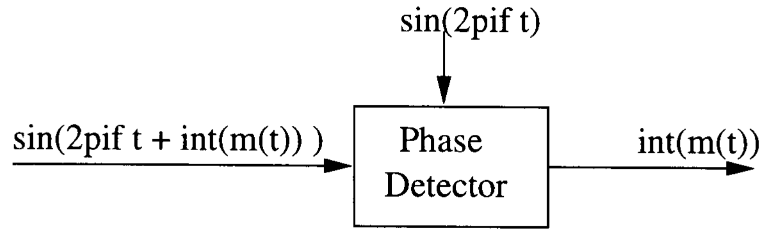

FIGURE 2.9: Phase Detector

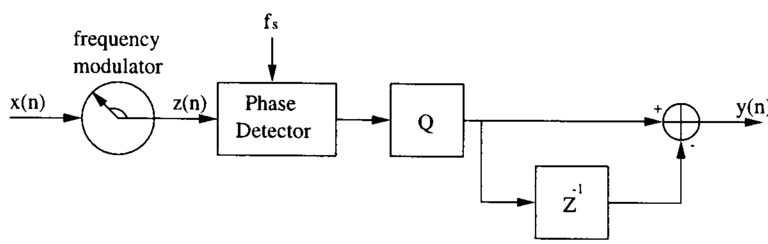

FIGURE 2.10: Hovin modulator demodulator

that y(n) is an estimate of x(n), and that z(n) is an FM signal, the balance of the circuit, from z(n) to y(n) must perform a demodulation.

The output of the phase detector is the phase of the FM signal. It operates by converting the delay between two different signals into a voltage or current. In the case of the Hovin demodulator, both input signals to the phase detector are limited using the sign() function, as was the case in the traditional ZCD. The output of the phase detector can be broken into two parts. An integer part, p_n, representing the number of FM zero crossings that were detected during the sampling period, and a fractional part ϕ_n, representing the phase difference between the previous rising FM edge and the sample signal edge, scaled by $1/2\pi$. [4]

The quantizer simply performs the floor function, as shown in figure 2.11. When the floor function is performed on the output of the phase detector, the fractional portion of its output, ϕ_n, will be removed, leaving only the integer part, p_n. Thus the quantization noise is equal to $-\phi_n$.

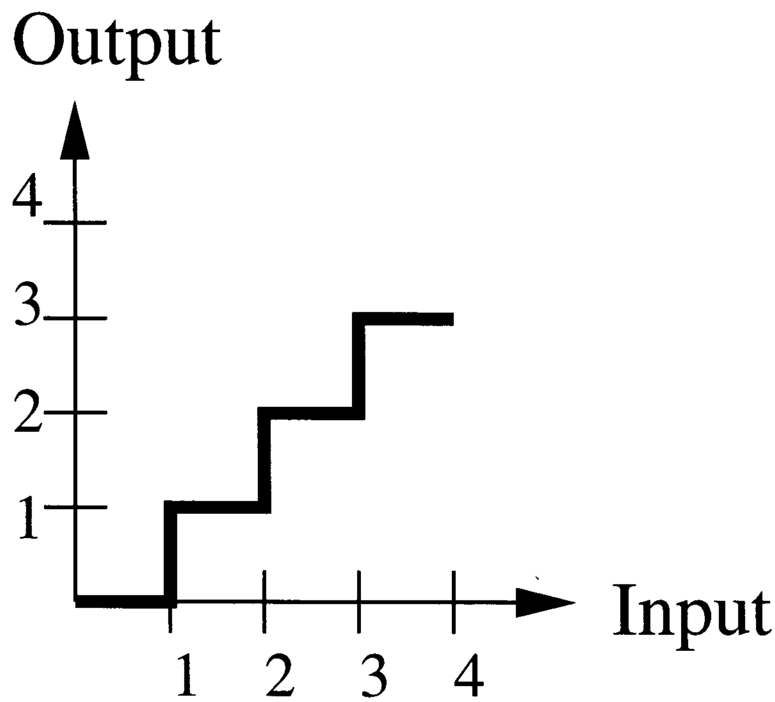

FIGURE 2.11: Graph of quantization function for Hovin Demodulator

Successive outputs of the quantizer block are subtracted from each other, which performs the differentiation on the signal that was integrated by the FM modulator and phase detector, equation 2.7. This extracts the original message signal, x(t). Since the quantization noise was added after the integration step, it is differentiated, which provides noise shaping. The circuit implementation of the Hovin demodulator is quite simple. If the phase detector, which counts the FM edges between the edges of its sampling clock signal, and the quantizer are combined, the fractional portion of the phase difference can be ignored. This is possible since the fractional portion of the period is lost during quantization. Therefore, the combination of the phase detector and quantizer reduces to counting the number of FM edges between sampling clock edges.

$$\frac{d}{dt}\int_{-\infty}^{t} x(\alpha)\delta alpha = x(t) \tag{2.7}$$

Recall that this signal represents an integrated version of the message signal. To extract the message the quantizer output must be differentiated. For discrete time signals that is accomplished by taking differences. Thus, successive outputs from the counter are subtracted, to extract the message signal. Therefore, the whole Hovin demodulator circuit consists of a counter and a subtracter. The counter must have enough range (bits) to be able to represent the difference between the maximum and minimum number of counts during one sampling interval.

A one bit version of the Hovin demodulator is possible, like in the traditional ZCD. For systems where the sampling frequency is more than twice the maximum FM frequency, the output of the counter will be constrained to one bit. In such a system, the counter is only a single flip-flop, and a one bit subtraction can be implemented by an exclusive-or gate, exactly the same as the one bit traditional method. This reduces the gate count for the demodulation portion of the receiver to 5 gates, a significant reduction.

Drawbacks to one bit demodulators include the fact that more filtering must be performed on the output to extract the signal, and the sampling frequency must be twice the maximum FM frequency, which will draw more power.

3. ANALYSIS

The goal of this thesis is to analyze the Hovin demodulator in terms of output signal quality and power consumption, which is closely related to gate count. Each of these two major categories has several subsections that are contributing factors to the analysis. The results of this analysis will be compared to the PLL demodulator and an ideal mathematical model of a demodulator. By the end of the chapter the reader should be able to determine which type of demodulator will perform best for a given set of design specifications.

The data presented in this section was generated using MATLAB to simulate the behavior of the circuits being compared. A MATLAB program was written to simulate each of the different demodulators. These simulators were then stimulated with similar inputs, and the outputs were compared. The target system is a wireless demodulator, so the input us usually a limited IF signal. For these simulations, the frequency of the IF wave is set to 100 kHz. In most cases the sampling frequency was set to 1 MHz, except when it was varied over a range to show what affect it has on system performance. The Δf is usually set to 5 kHz. The message in the simulations is a 1kHz sin wave which is easy to compare to the original, and therefore determine noise performance.

3.1. Output Signal Quality

Output signal quality is probably the most obvious means of comparison for any type of communications circuit. The main goal in any communications system is to have an accurate representation of the message signal at the system output. To reach this goal, each block of the system needs to maintain the best possible level of signal integrity.

In FM demodulators there are several different factors that deteriorate the output signal quality. First for a digital demodulator, like Hovin and the traditional zero crossing demodulator, quantization noise is a big factor in the output signal quality. In the zero crossing demodulators, quantization noise is going to be added by the counters present in both the Hovin and the traditional methods. In the traditional method, the noise arises from the fact that the pulse output of the zero crossing detector does not exactly coincide with the zero crossing, but is a close approximation. In Hovin's method, the noise arises from the fact that the quantizer removes the fractional portion of the phase detector output.

Second, thermal noise present at the input to the demodulator, produces noise in the output signal. For FM demodulators the relationship between the input thermal and output noise is non-linear. To characterize this effect it is useful to compare the carrier-to-noise ratio (CNR) at the input to the signal-to-noise ratio (SNR) at the output.

Third, there is often harmonic distortion present in the output signal, since FM demodulation is a non-linear process. This distortion is due to odd and even order harmonics of the input tone being present in the output signal, and is related to the input signal power and frequency.

The parameters that we can control to affect these distortions are Δf, and the sampling rate. A system designer using zero crossing demodulators as receivers can adjust the Δf of the system to improve demodulator performance. In the Hovin and traditional zero crossing demodulators, the sampling rate can be adjusted at the receiving end to improve output signal quality. Making adjustments to these parameters will have trade-offs, as with many engineering applications.

3.1.1. Quantization Noise

Signal to Quantization Noise Ratio (SQNR) is a way to measure how much noise is added to the signal by the quantization step. Quantization noise will only be a factor in digital demodulators, since it is added by the quantizer that converts the analog signal into its digital representation.

In the case of the traditional zero crossing demodulator and the Hovin demodulator, the quantization noise added to the signal is shaped. The noise power is pushed up to higher frequencies, much like the noise in a sigma delta modulator. In order for the noise to be pushed out of the band of interest, the ZC demodulator must sample at a rate that is much higher than the Nyquist rate. Thus, most of the quantization noise in the demodulator can be easily filtered. Raising the sampling rate also improves resolution. Since the clock is ticking faster, we have a more accurate representation of when the zero crossing actually occurred. The higher the sampling rate, the farther the noise is pushed up, leaving less noise in the band of interest. Of course, a higher sampling rate means more power consumption.

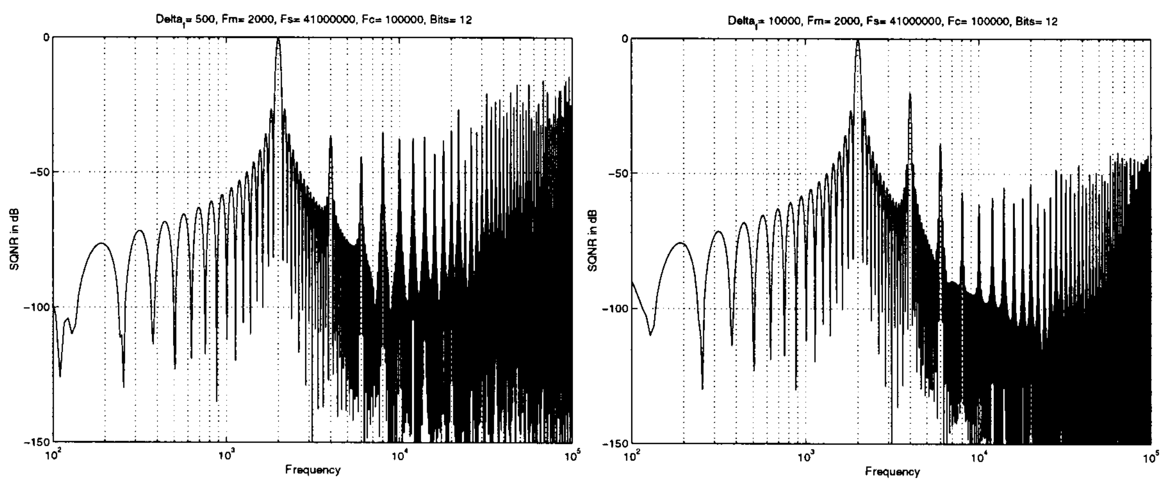

FIGURE 3.1: Noise shaping for 12 bit ZC demodulator, 2 kHz tone, 2^{15} point fft, 41 MHz sampling rate. Left: Δf = 500, Right: Δf = 10k

Figures 3.1 and 3.2 reveal the noise shaping in the traditional and Hovin demodulators for several values of Δf. Figure 3.1 shows the noise shaping associated with a traditional zero crossing demodulator. For frequencies above 10 kHz, noise shaping can be observed. This noise shaping has a slope of about 40 dB per decade, which is consistent with expectations. Below the 10 kHz mark spectral leakage begins to dominate the spectrum of the demodulated signal. This is due to the problem with data output from

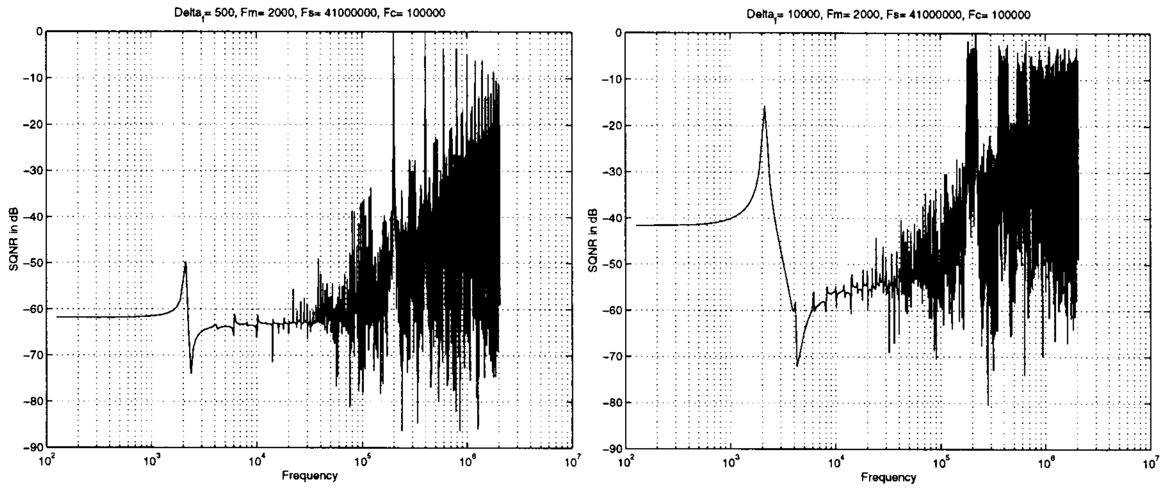

FIGURE 3.2: Noise shaping for 12 bit Hovin demodulator, 2 kHz tone, 2^{15} point fft, 41 MHz sampling rate. Left: Δf = 500, Right: Δf = 10k

the counter occurring at zero crossings of the IF signal. The true frequency of the data signal is distorted since its samples are arriving at a changing rate.

Figure 3.1 also demonstrates how a higher value of Δf increases demodulated signal power. As Δf gets larger, the peak value of the signal power grows taller in relation to the surrounding noise. This is because at higher values of Δf, there is a greater difference between the maximum frequency and the minimum frequency present in the FM signal. A greater difference allows the zero crossing demodulator to demodulate with greater resolution.

The noise shaping produced by the Hovin style demodulator, Figure 3.2, is much more pronounced. The noise is not dominated by harmonics, as in the traditional method, but is mostly just sampling noise. The shape, a 30 dB per decade increase in noise, is just what was expected.

Figure 3.2 demonstrates the received signal powers dependence on Δf, just as figure 3.1 did. As Δf increases, the height of the spike of the signal power increases in relation to the noise floor.

Gains from increasing the Δf are only effective up to a certain point. If Δf gets too low, the zero crossing demodulators cannot detect any difference between the maximum and minimum periods of the FM signal. If this is the case, the output of the demodulator will be a constant value. A larger modulation constant will make it easier to discern between a short period and a long period of the input FM signal. The larger constant will also mean the number of different values the period can assume will be larger. After a certain point, no more advantage is gained by making Δf larger. In the traditional zero crossing demodulator, this is partly due to the distortion caused by a non- constant data rate, described above. In fact, if the Δf gets too large, the distortion degrades the signal output significantly, as seen in figure 3.3.

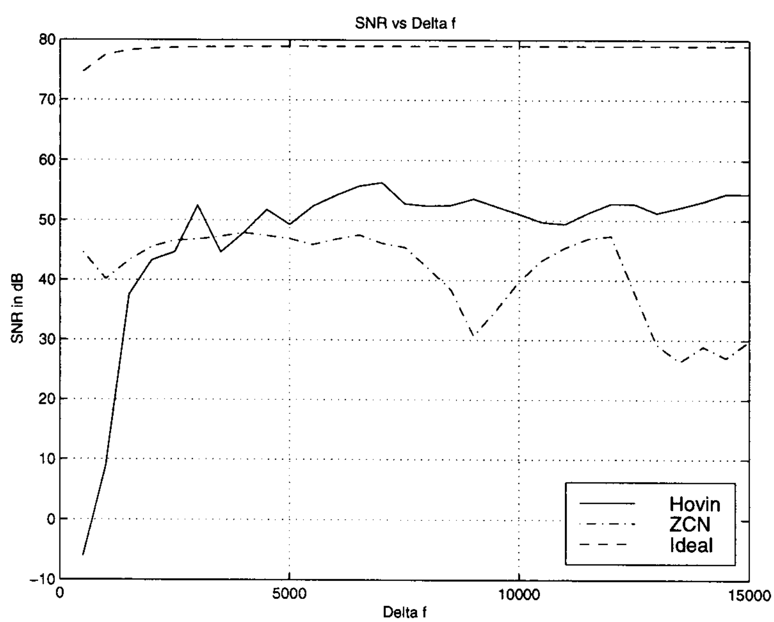

FIGURE 3.3: SNR vs Δf for Ideal, ZCN and Hovin demodulators

The digital sampling frequency of a zero crossing demodulator is also a contributor to the amount of quantization noise that is present in the system. In the case of the traditional demodulator, sampling frequency is a double edged sword. Increasing sampling frequency will improve resolution and increase noise shaping benefits. It will also increase power consumption and distortion. Increasing the sampling frequency divides the input FM wave into smaller chunks, which are then counted. Since there are more chunks to count, smaller differences in frequency can be detected, in the same way that increasing Δf increases resolution. Noise shaping is also dependent on sampling frequency, a higher sampling rate drives the quantization noise up to a higher frequency band. The distortion due to increased sampling frequency is identical to that caused by a larger Δf. Increasing the sampling frequency also means that more bits are needed to represent the count values, increasing the gate count. A higher gate count coupled with the fact that the gates are toggling faster increases power consumption

Figure 3.4 shows just how much the distortion due to the variable data rate degrades the traditional zero crossing demodulators output. For the one bit case, ZC1, the output SNR rises quickly and then is nearly constant. This is because the one bit ZC is not affected by the variable data rate distortion. Comparing figures 3.4 and 3.5, shows the benefits, and pitfalls, of a high sampling rate coupled with a large Δf. Figures 3.4 and 3.5 were produced from the ensemble average of several runs of the simulator. Each run used a data signal that started with a random phase.

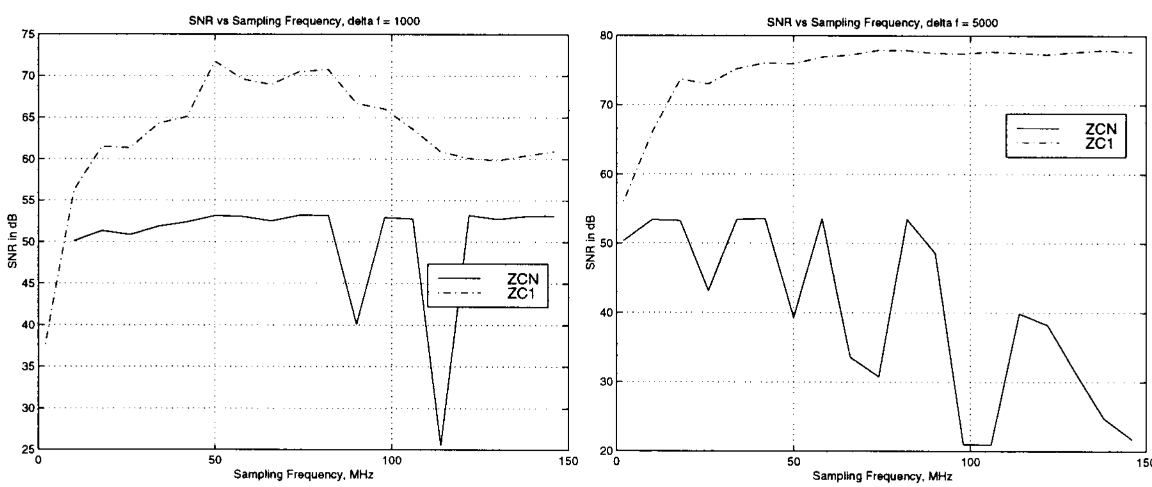

FIGURE 3.4: Narrow Band FM SNR vs Sampling frequency for zero crossing demodulators, Δf = 1000, 5000; Fm = 10k

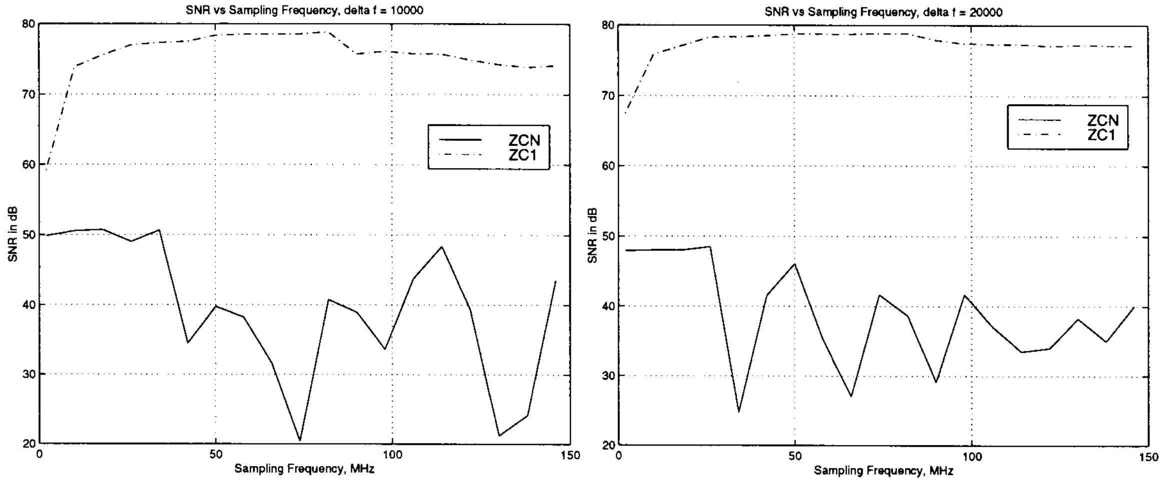

FIGURE 3.5: Wide Band FM SNR vs Sampling frequency for zero crossing demodulators, Δf = 10k, 20k; Fm = 10k

3.1.2. SNR vs CNR

SNR vs CNR is a measure of how much noise in the transmission channel the system can tolerate before the demodulator cannot demodulate [6]. Since FM systems are immune to a certain amount of noise, they can tolerate a moderate amount with almost no signal degradation. When the noise gets too powerful, the output signal quality degrades quickly.

The SNR vs CNR for the ZC demodulator depends on several things. It is most strongly a function of the modulation constant, Δf. The SNR vs CNR will also depend on the quality of the amplifier and limiter in the IF stage. If they do not work together to produce a high-quality square wave from the input, the SNR will suffer.

Figures 3.6 and 3.7 show a shape that is characteristic of FM demodulators. The curve is linear for higher CNR's and then degrades quickly when the threshold is reached. This curve can be used to determine a minimum CNR required for correct operation of the circuit. Armed with the minimum CNR required for demodulation, the system designer can choose the rest of the components in the system, so that the CNR will be met.

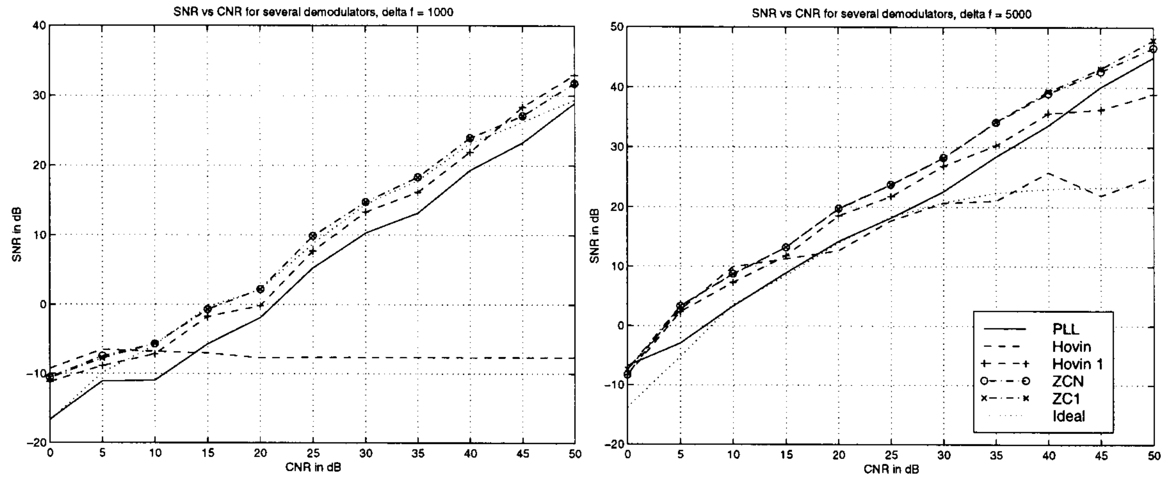

FIGURE 3.6: Narrow Band FM SNR vs CNR, Δf = 1000, 5000; Fm = 10k, Fs = 10 MHz

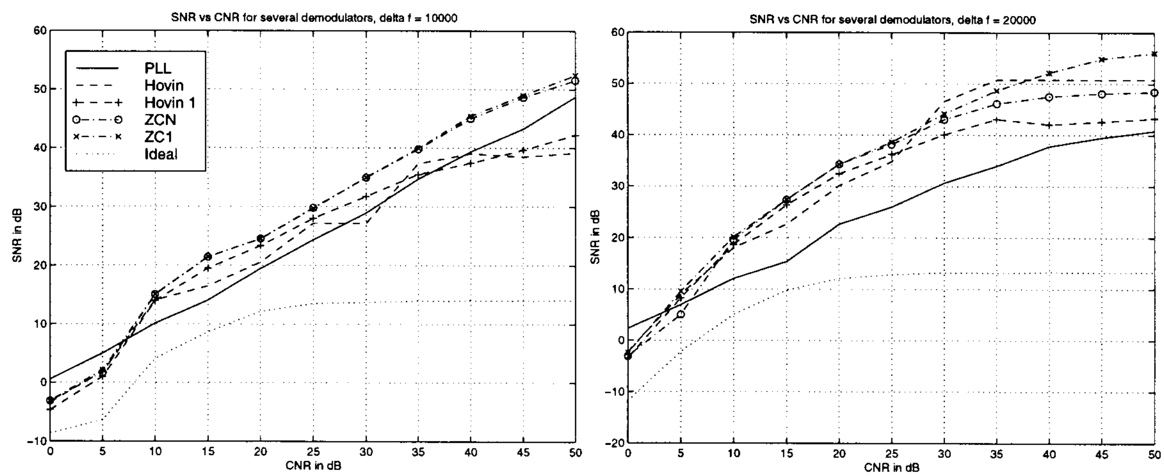

FIGURE 3.7: Wide Band SNR vs CNR, Δf = 10k, 20k; Fm = 10k, Fs = 10 MHz

Figure 3.6 shows how the Hovin demodulator does not perform well at all when the Δf is too small. In the graph on the left the Hovin demodulator gets no improvement with more signal power. The graphs in Figure 3.7 show an unusual shape for the ideal demodulator curve. This poor performance seems to be a problem when demodulating noisy wide band FM signals using MATLAB's demod routine.

The most interesting features of this graph are the large step in the Hovin demodulators curve, and where each curve levels off. The step in the Hovin curve indicates that it does much better than any of the other three types of demodulator for that certain range. The two zero crossing based demodulators also level off at a lower value than the ideal and PLL based demodulators. This means that the maximum SNR output for the zero crossing methods is lower than that of the other types of demodulator.

This higher SNR output for the PLL makes it more attractive for systems that demand a very high quality output. There are several drawbacks to using this kind of demodulator. First the curve also shows a linear relationship between input and output SNR. If the input CNR degrades for a short period of time, the output SNR will be affected, whereas in a zero crossing style demodulator, as long as the input does not drop below the threshold value, the output signal quality is about constant.

Δf	Ideal	PLL	Hovin	ZC
500	50 dB	75 dB	-8 dB	45 dB
1000	50 dB	76 dB	20 dB	40 dB
5000	50 dB	78 dB	50 dB	47 dB
10000	50 dB	78 dB	52 dB	40 dB

TABLE 3.1: Comparison of SNR vs Δf for several Demodulators. Each demodulator optimized once.

3.1.3. Harmonic Distortion

Harmonic distortion poses a different problem than the previous two forms of signal degradation. Quantization noise and CNR are caused by outside sources. Steps can be taken by the designer to minimize their harmful affects. Harmonic distortion is caused by the message signal itself as it passes through the modulator and demodulator.

Two different types of harmonic analysis were performed on the group of demodulators. The first analysis was to determine Total Harmonic Distortion for each of the demodulators. Total Harmonic Distortion (THD) is a measure of how much distortion the demodulator adds to the signal. It is a sum of the signal present at each harmonic of a tone that has been demodulated. The percentage of THD is calculated using a single tone test. A single tone is modulated and then demodulated by the demodulator being analyzed. The resulting demodulated signal is used in equation 3.1, which sums the power present at each integer multiple of the fundamental frequency of a test tone. [7]

$$THD(\%) = \frac{\sqrt{\sum_{n=2}^{\infty} V_n^2}}{V_1} \times 100 \tag{3.1}$$

THD calculations were performed on each of the demodulators for varying values of Δf. The results are shown in Figure 3.8. The figure shows that for the Hovin and traditional ZCD, the harmonic distortion is low for smaller values of Δf, and then drops quickly as Δf gets larger. This is due to the fact that it is easier for a zero crossing type demodulator to sense the difference in frequencies if they are farther apart. The THD for the PLL demodulator starts off low, and then gets larger very quickly. This is due to the fact that a PLL can track a slowly varying frequency much more easily than it can track one that varies quickly. It should be noted that a different PLL could be designed that will track frequencies that vary more quickly, if one was needed. The PLL information is included here for comparison purposes [8]. Also note that for very low values of Δf, the zero crossing and Hovin demodulators will not demodulate the signal at all, which is why there are no data points for those two demodulators at low values of Δf.

The second type of harmonic analysis performed on the demodulators involves a two tone test. Two pass band tones are modulated and then demodulated by each demodulator. This test shows what happens when several tones are passed through the demodulator at the same time, and is called an Inter-modulation Distortion test (IMD). It is called Inter-modulation distortion because it is only present when two or more input frequencies are present.

IMD distortion has a different character than THD, and causes different problems. The IMD is due to the difference frequencies caused by non-linearities in the demodulator,

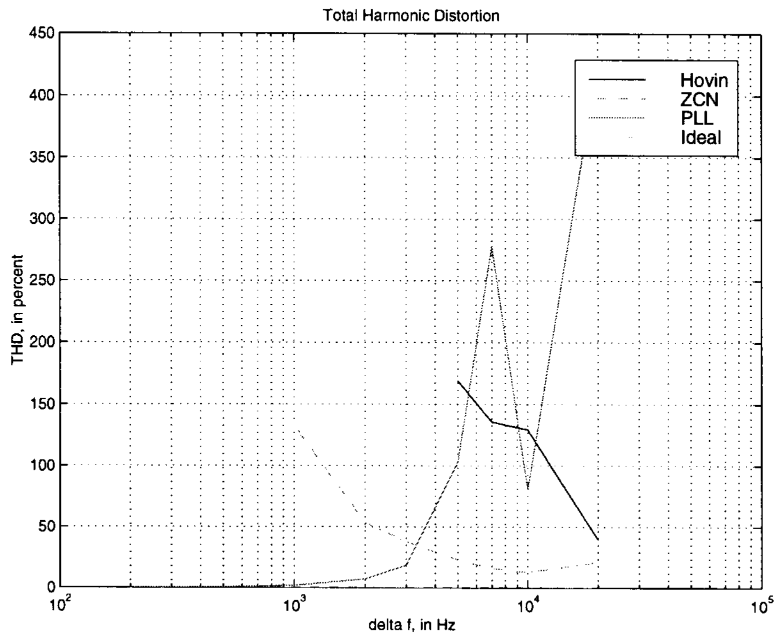

FIGURE 3.8: Total Harmonic Distortion for several demodulators

which fall very close to the original frequencies. These harmonics are not easily removed, since they may fall well within the band of interest. The THD harmonics, however, are multiples of the frequencies present in the signal, and most of them will fall outside the band of interest.

Two different two tone tests were performed. The first test explores harmonic distortion due to input signal level. To do this it varies the modulation constant Δf over a range and checks the signal power at each of the IMD frequencies. For both demodulators this produces the results that are expected, a steadily decaying level, figure 3.9. The harmonic level decreases as the modulation constant increases because the demodulator is able to get a more accurate measurement of the data signal level. When the modulation constant is high, there is a greater difference between the lowest frequency present in the IF signal and the highest frequency. Therefore there will be a larger difference in the counts during the shortest period and the longest period.

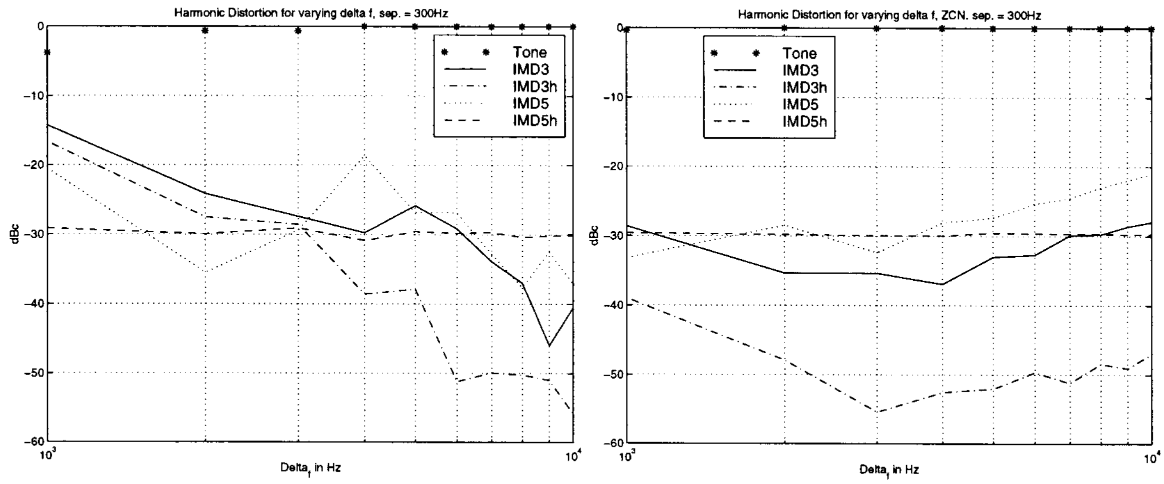

FIGURE 3.9: Harmonic distortion due to input signal level. Left: Hovin demodulator, Right: ZC demodulator

Next, harmonic distortion due to separation of the two input data tones is explored. To do this the tone separation is varied and the output level at each of the harmonic frequencies is checked.

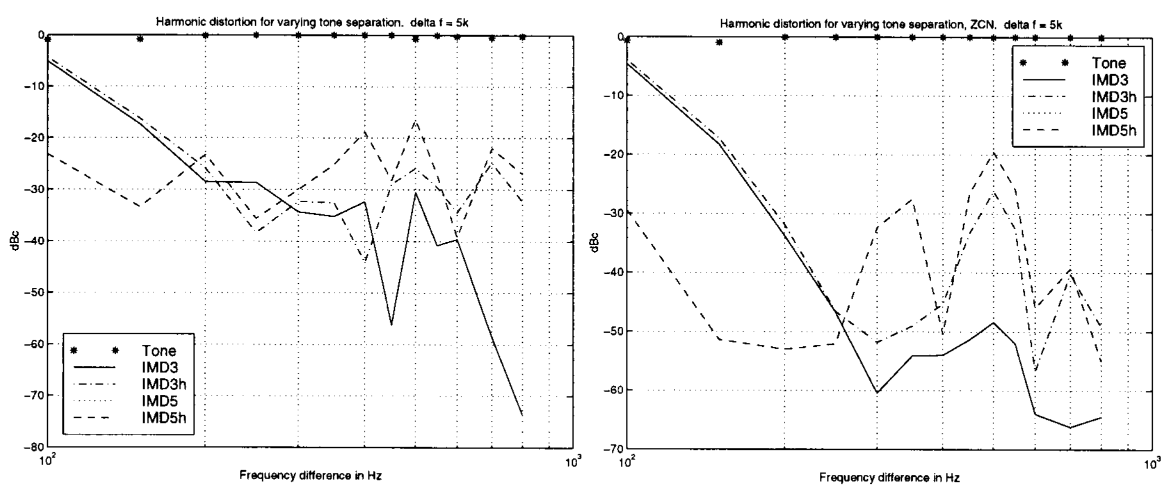

FIGURE 3.10: Harmonic distortion due to tone separation. Left: Hovin, Right: ZC

Three types of signal degradation have been examined. From the analysis it can be seen that the amount and type of signal degradation is a factor of three choices that can be made by the designer. The first is the type of demodulator used. This will have a big effect on whether or not quantization noise is an issue at all. The second choice is the Δf of the system. It has been demonstrated that for each type of demodulator, different factors play a role in determining the best choice for Δf. The third choice that the designer can make is the sampling frequency that the demodulator will use. Sampling frequency will affect how much quantization noise is added to the output signal, and in some cases, how much harmonic distortion is present.

3.2. Power Consumption

In a wireless receiver, power is provided by a battery. To make the battery last as long as possible, power usage must be minimized. In certain wireless applications this factor can override many other considerations, including signal quality, and circuit complexity.

Gate count is closely related to power consumption in a digital synchronous system. It also is a rough determination of the size of the chip required to implement a circuit. In a digital CMOS IC, size is also inversely proportional to yield. A lower yield results in higher per unit costs, lowering profits.

The Hovin demodulator has a much lower gate count than a ZC demodulator for the same performance at the output. The Hovin demodulator's gates are also being clocked at a much lower rate than the ZC gates are.

The Hovin demodulator uses roughly four gates per counter bit in it's first stage. This means that if the counter is N bits, $4 \times N$ gates are needed for the counter. These gates will be running at approximately the same speed as the IF signal. To store the counter output when it is needed, a register is needed to hold the data while the counter is being reset. This register will take another $4 \times N$ gates. This register will only be loaded once for every zero crossing in f_h. One gate is needed to generate a load / reset pulse for the register and the counter. This gate will also run at the f_h speed. If a comb filter is used to filter the output data, it must be N bits wide, and M stages long. The value of M

is determined by f_h and by the desired data rate at the output. The comb filter consists of two parts, a delay line, and an adder. The delay line is merely a series of M registers. The adder adds all of the delayed values together. The delay line will require $4N(M-1)$ gates. These gates will be clocked at f_h. The adder can get a bit more complex. If the desired data rate is very high, a simple ripple carry counter may not be fast enough to add all M stages together in the amount of time given. If so, a faster adder will have to be used. The ripple carry adder is the smallest, so it consumes the least power. A faster adder, such as Ladner-Fischer, or Parallel prefix adder would be much more complex and would draw more power than the ripple carry. The adder will run at $M \times f_h$ since it has to add all of the stages together between zero crossings.

Approximate values of N and M can be calculated from equations 3.2 and 3.3. The ZC demodulator has more gates, and most of them are running at a higher clock rate. The input stage is 2 flip-flops and an XOR gate, running at the IF rate. The counter in the ZC demodulator is running at the sampling rate, f_s, which is at least 10 times higher than the IF rate. The counter needs to have N bits, which will require $4N$ gates. The ZC demodulator requires a register to store the data, just as the Hovin does, which means another $4N$ gates, running at the IF. Following the register, the ZC demodulator requires an interpolation circuit, in order to synchronize the output of the demodulator with a regular clock. The interpolater can be realized using a comb filter where the IF signal clocks the delay line, and the output is fed into another comb filter. The second comb filter is clocked with a synchronous signal that is generated so that the data will be output at the correct rate.

A PLL is an analog device, so it doesn't have gates in the same sense as a digital device. Power consumption in a PLL will depend on other factors.

$$N = \left\lceil log_2 \frac{f_{IF}}{f_h} \right\rceil \tag{3.2}$$

$$M = \left\lceil \frac{f_h}{f_{data}} \right\rceil \tag{3.3}$$

$$N \approx log_2 \left\lceil \frac{f_s}{f_{IF}} \right\rceil \quad (3.4)$$

	SNR	Bits	$Gates \times f$
Hovin	58.3 dB	5	4.3 Million
ZCN	61.3 dB	9	2.9 Billion

TABLE 3.2: Power comparison for comparable SNR performance. Δf = 5k, f_c = 100k, f_m = 5k

Table 3.2 shows a comparison of the Hovin style demodulator with the ZCN demodulator. The two demodulators were designed to demodulate the same signal. Each demodulator was optimized to produce an output SNR of about 50 dB. The table shows the relative power consumption of each of the demodulator circuits. The last column in the table shows the number of gates at a certain clock frequency times that clock frequency. This number is linearly related to the power that that circuit consumes. The table shows that for similar noise performance, the Hovin demodulator consumes three orders of magnitude less power than the N bit Zero Crossing demodulator. This makes the Hovin demodulator much more attractive for wireless applications, if the Δf of the transmit system is large enough.

4. CONCLUSIONS

The purpose of this thesis was to evaluate a new type of zero crossing demodulator, the Hovin demodulator, in terms of output signal quality, power consumption, and gate count. The Hovin demodulator was compared to traditional zero crossing demodulator, a PLL demodulator and an ideal demodulator.

The analysis section shows that both types of zero crossing demodulators can perform just as well as a PLL system, and in some cases almost as well as an ideal demodulator. Figures 3.6 and 3.7 demonstrate that the zero crossing demodulators can perform well in a noisy environment. Analysis of distortion and quantization noise shows that both can be reduced with proper choice of circuit parameters. Therefore it has been demonstrated that both the Hovin ZCD and the traditional ZCD can perform just as well as demodulators currently in use.

One of the main advantages of the zero-crossing demodulator over an analog demodulator is that it performs analog to digital conversion in the same step. If a PLL was being used, it would have to be followed by some sort of ADC to make the data useful. In some cases this ADC would be as complex and draw as much power as the demodulator itself, and increases the circuit complexity and size as well. Using a zero crossing demodulator saves having to use an additional ADC. This way the zero crossing demodulator also eliminates another source of error and distortion, the ADC itself.

The zero-crossing demodulators are entirely digital. One of the advantages of being all digital is that the demodulator can be improved by improving the CMOS technology that is used to implement the demodulator without changing the design of the demodulator itself. One design could be updated to many generations of new technologies that are capable of running at higher clock speeds without doing any re-design work. With maximum clock speeds increasing at the current rate, a practical, all-digital demodulator that does not use an intermediate frequency stage, could become a reality in the next few years.

Being an analog to digital converter of sorts, the zero crossing demodulators do add a certain amount of quantization noise. However the amount of noise that is added, and the band of frequencies that the noise occupies can be controlled by the designer. This way, a system can be designed that will meet almost any noise performance requirements.

The regular ZC demodulators biggest advantage is it's robustness. A one bit ZC demodulator can demodulate just about any FM signal as long as the sampling rate is high enough. So one circuit could be built and different clocks fed into it to demodulate whatever is presented to the circuit. This circuit is very easy to "tune". The clock frequency can merely be increased until the desired SNR output is achieved. The disadvantage to this scenario is, of course, power consumption. As the clock frequency goes up, so does the power drawn.

The Hovin demodulators biggest advantage is that it consumes much less power than it's cousin the traditional ZCD, since it can be implemented using far fewer gates. The gates in the Hovin implementation will also be running at a slower rate than those in the traditional ZCD. As can be seen from table 3.2, the Hovin style demodulator power consumption is several orders of magnitude lower than that of a traditional ZCD. The Hovin demodulator does not require an interpolation stage to produce synchronous data, like the traditional ZCD does. The main draw-back is that it needs a certain amount of Δf in order to demodulate at all. In some systems this could hamper the Hovin ZCDs usefulness.

BIBLIOGRAPHY

1. Jack R. Smith. *Modern Communications Circuits.* McGraw Hill, 2nd edition, 1998.
2. Simon Haykin. *Communication Systems.* Wiley, 3rd edition, 1994.
3. Mats Hovin, Trond Saether, Dag T. Wisland, and Tor Sverre Lande. A narrow-band delta-sigma frequency-to-digital converter. *IEEE*, ?(?), March 1997.
4. Mats Hovin, Alf Olsen, Tor Sverre Lande, and Chris Tomazou. Delta-sigma modulators using frequency-modulated intermediate values. *IEEE Journal of Solid-State Circuits*, 32(1), January 1997.
5. Mats Hovin. Novel delta-sigma modulators using frequency modulated intermediate values. Master's thesis, University of Oslo, 1995.
6. Herbert Taub and Donald Schilling. *Principles of Communication Systems.* McGraw Hill, 1971.
7. Leon W. Couch II. *Digital and Analog Communication Systems.* Prentice Hall, 5th edition, 1990.
8. Richard Gitlin, Jeremiah F. Hayes, and Stephen B. Weinstein. *Data Communications Principles.* Plenum Press, 1992.

Appendices

A SECOND ORDER IMPLEMENTATIONS

Both the ZC and the Hovin demodulators may be improved by using second order implementations. Second order implementations would give greater noise shaping, with a 40 dB per decade roll-off. This steeper roll-off would allow for more room in the passband without increasing sampling rate, and for filters with a gentler roll-off to filter out the high frequency noise.

To implement a second order zero crossing system, a way must be found to quantize the fractional part of the sampling period where the zero crossing occurs. For example, if the zero crossing occurs exactly half way between two clock edges, we need to output some value that represents half. If it occurs in the first 10 percent of the period, we need to output a value corresponding to 10 percent. There are several different ways to do this. One way is to use a second very high speed clock. It would divide the period between the sampling clock into much smaller intervals. The main problem with this approach is that if a much faster clock could be implemented, often it would already be the sampling clock. A higher clock speed also consumes more power. A second method is to charge a capacitor during each interval. The capacitor charges at a nearly linear rate during the period, and when the zero crossing occurs, the capacitor voltage is compared with a known voltage level. This approach requires a much more complex circuit, and some sort of analog comparator or analog to digital converter.

B DESIGN TOOL

The Tool takes design specifications in and spits out the different constants and frequencies that are needed to create the demodulator. The Tool takes F_{if}, SNR, F_{data} as inputs and calculates F_h, M, N, and P for the circuit, where F_h is the Hovin sampling frequency, M is the number of stages required in the comb filter, N is the number of bits needed in the counter, and P is the estimated power consumption.

The first thing it has to do is determine is what type of demodulator to use to best meet the needs of the designer. Then it has to determine what sampling frequency, F_h, or F_s, to use. From this and the IF frequency, it can determine how large the counter must be to handle the highest value using

$$N = \lceil log_2(F_{if}/F_h) \rceil \tag{2.1}$$

Then using F_h and F_{data} it can be determined how many comb filter stages are required to filter the data.

$$M = \lceil F_h/F_{data} \rceil \tag{2.2}$$

Using these two numbers, it can then calculate how many gates are needed to build the circuit, and from that it can estimate power consumption. The following formula is being used for power consumption calculations:

$$P = CV^2 f(DC) \tag{2.3}$$

where P is the power, C is the gate capacitance of the gate that this one is driving, V is the power supply voltage, f is the frequency at which the gate switches, and DC is an optional duty cycle factor. This equation is used to calculate the power consumed by a single gate, so it must be multiplied buy the number of gates at the frequency to determine sub-totals which can then be summed. For example:

$$P = CV^2(2NF_{if} + 2NMF_{hovin}) \tag{2.4}$$