AN ARCTANGENT TYPE WIDEBAND PM/FM DEMODULATOR

WITH IMPROVED PERFORMANCES

Noël Boutin and Hatem Kallel
Université de Sherbrooke
Faculté des sciences appliquées
Département de génie électrique
Sherbrooke, Québec, Canada - J1K 2R1

ABSTRACT

Based on the idea that a PM or FM signal can be seen as a QAM signal some PM/FM demodulators using a QAM demodulator are presented. Among them, the arctangent type is studied in more detail and simple modifications are proposed to improve its performances.

1. INTRODUCTION

Owing to their superior performances over AM, phase (PM) and frequency (FM) modulated signals are currently used in many analog and digital communication systems. The well known method to demodulate those PM/FM signals is the superheterodyne receiver equiped with an intermediate frequency (IF) amplifier-filter combination followed by a limiter-discriminator circuit. That architecture is however not well suited with modern trends such as VLSI integration and digital signal processing (DSP). Alternative architectures can be found if one understand the true nature of PM/FM signals.

2. Visual Interpretation of PM/FM signals

Let us consider the following constant envelope phase or frequency modulated sinusoïdal carrier of nominal angular frequency ω_c

$$s(t) = 2\cos[\omega_c t + g(t)] \tag{1}$$

The modulating signal g(t) is related to the message m(t) to be transmitted in the following way

$$\text{PM case} \qquad g(t) = K_p m(t) \qquad \text{rad} \tag{2}$$

$$\text{FM case} \qquad g(t) = K_\omega \int_0^t m(t)\,dt \qquad \text{rad} \tag{3}$$

where K_p and K_ω are the conversion factors of the modulator respectively expressed in radian/volt and radian/sec/volt, with m(t) in volts.

The signal s(t) given in (1) can also be expressed as

$$s(t) = \cos[g(t)] \cdot 2\cos\omega_c t - \sin[g(t)] \cdot 2\sin\omega_c t \tag{4}$$

The main issue now is to recognized that the real RF signal s(t) given in (4) is the result of the addition of a complex baseband signal c(t) shifted in frequency by an amount $+\omega_c$ with the conjugate version $c^*(t)$ of that complex baseband signal shifted in frequency by an amount $-\omega_c$. That is,

$$s(t) = c(t) \cdot e^{j\omega_c t} + c^*(t) \cdot e^{-j\omega_c t} \tag{5}$$

where

$$c(t) = \cos[g(t)] + j\sin[g(t)] \tag{6}$$

Contrary to the usual practice, is is clear from (5) that the focus must be put on the complex baseband signal c(t) rather than on the RF signal s(t) which is only a frequency shifted real version of c(t).

(6) can also be expressed as

$$c(t) = e^{jg(t)} \tag{7}$$

In spite of the fact that this representation is known since many years [3], it seems that, until recently [4,5], nobody has presented a clear interpretation of the meaning of (7). What (7) tells us is that the generation of a PM or an FM signal, by whatever means, is tantamount to non-linearly process a real modulating baseband signal g(t) in such a way that the resultant baseband processed signal c(t) is complex. In other words, the two-dimensional (real amplitude vs time) signal g(t) is converted into a three-dimensional (complex amplitude vs time) signal c(t). And, according to (7), that processing is done **by enrolling g(t) on the surface of a unity radius cylinder.**

3. PM/FM demodulation

From (4), it turns out that a PM/FM signal could be seen as a Quadrature Amplitude Modulated (QAM) signal and demodulated as such with a QAM demodulator. Figure 1 illustrates that particular QAM communication system.

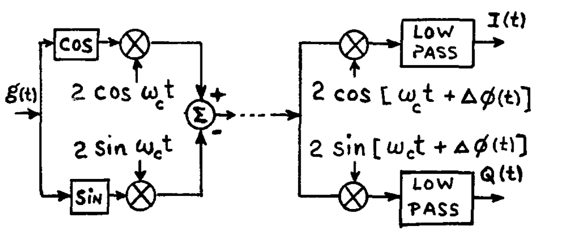

Figure 1: PM/FM communication system seen as a QAM system.

The signal present at the output of the QAM demodulator is not the original real baseband signal g(t) but, at most, the orthogonal projections I(t) and Q(t) of the complex baseband signal c(t) which is a rolled up version of the original real baseband signal g(t).

CH2819-1/90/0000-0460$01.00©1991 IEEE

Assuming for the moment a perfect local oscillator synchronization ($\Delta\phi(t) = 0$), $I(t) = \cos[g(t)]$ and $Q(t) = \sin[g(t)]$. The baseband real signal g(t) could therefore be regenerated, as shown in figure 2, by taking the ratio of Q(t) over I(t) followed by the arctangent operator.

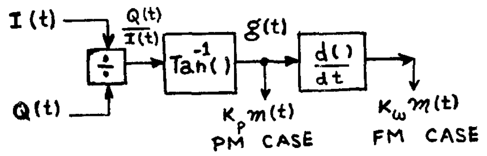

Figure 2 Conceptual baseband unwrapper.

That conceptual baseband unwrapper presents two drawbacks which preclude its direct materialization. First, each zero crossing of I(t) forces the divider output to saturate. Second, the signal present at the output of the arctangent operator is limited to the principal values of $\pm\pi$ since I(t) and Q(t) vary between ± 1. That signal is therefore equal to g(t) only in the narrowband situation where the peak amplitude of g(t) is restricted at the transmitter to ± 1.

3.1 Quadricorrelator

The aforementioned drawbacks have been bypassed in the past by making use of the following identity

$$\frac{d}{dt}\left[\arctan\left(\frac{Q(t)}{I(t)}\right)\right] = \frac{I(t)\cdot\frac{dQ(t)}{dt} - Q(t)\cdot\frac{dI(t)}{dt}}{I^2(t) + Q^2(t)} \tag{8}$$

Park [6] seems to be the first one to propose a baseband processor based on that relation. A quite similar arrangement has also been proposed by Denenberg [7]. In situations where a bandpass-limiter or other forms of normalization are used, the materialization is simplified because the denominator is a constant [8-11]. The resultant demodulator has been called a "balanced quadricorrelator" by Gardner [12] owing to its similitude with a basic earlier frequency difference detector proposed by Shaeffer [13] and later described in some detail and named "quadricorrelator" by Richman [14].

In situation where a digital implementation is considered [15-16], the differential terms in (8) are often approximated by

$$\frac{dQ(t)}{dt} = \frac{Q(t) - Q(t-\tau)}{\tau} \qquad \frac{dI(t)}{dt} = \frac{I(t) - I(t-\tau)}{\tau} \tag{9}$$

where τ is an arbitrary small time delay.

Substituting (9) in (8), one gets

$$\frac{d}{dt} = \left[\arctan\left(\frac{Q(t)}{I(t)}\right)\right] \approx \frac{-1}{\tau}\cdot\frac{I(t)\,Q(t-\tau) - Q(t)\,I(t-\tau)}{I^2(t) + Q^2(t)} \tag{10}$$

One should recognize that all the demodulators based on relation (8) are intrinsically FM demodulators as opposed to the PM demodulator shown in figure 2. Furthermore, the approximation given in (10) does not yield to a linear characteristic but a sinusoïdal one.

3.2 Extended range arctangent type demodulators

Some years ago, Hagiwara and Nakagawa [1-2] reconsidered the baseband unwrapper shown in figure 2 and proposed the modified version of figure 3 which bypass the two drawbacks of the earlier, that is the saturation of the divider and the limited range of phase modulation.

We conjecture that better performances could be obtained if one relocalizes the time differentiator from the input to the output of the range extender. Doing so, less impulsive noise would be present at the input of the range extender and less jump detection errors would occur. In principle, that simple modification could improve performances at low input signal-to-noise ratio (S/N)i.

To validate our assumption, we simulated both systems using exactly the same environment. The modulating signal g(t) was chosen as $\beta \sin \omega_m t$, the modulation index β being a variable parameter.

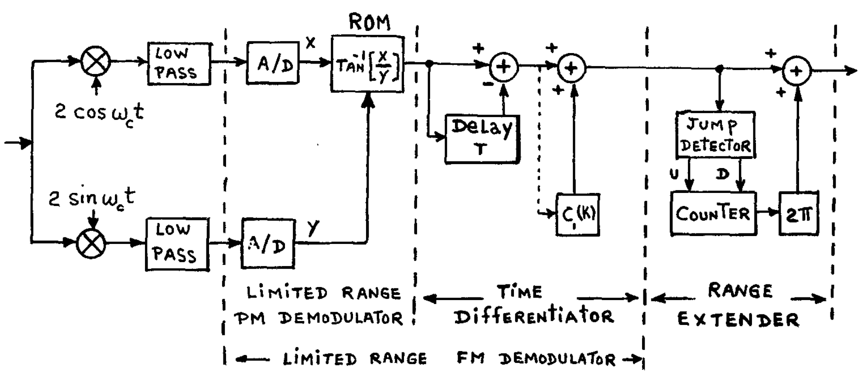

Figure 3: The "extended range tangent type" demodulator proposed by Hagiwara and Nakagawa.

Additive white Gaussien noise was added to the transmitted signal s(t). An input Butterworth 3^{rd} order band-pass filter was inserted ahead of the QAM demodulator. Its -3dB bandwidth was fixed to 2.4 $(1+\beta)\omega_m$. The two post-demodulation low-pass filters were 3^{rd} order Butterworth low-pass with -3dB cut-off frequency equal to 1.5 $(1+\beta)\omega_m$. Finally, the demodulated signal g(t) was filtered with a 2^{nd} order lowpass Butterworth filter having a -3dB cut-off frequency equal to ω_m.

The results of our simulation appear in figure 4.

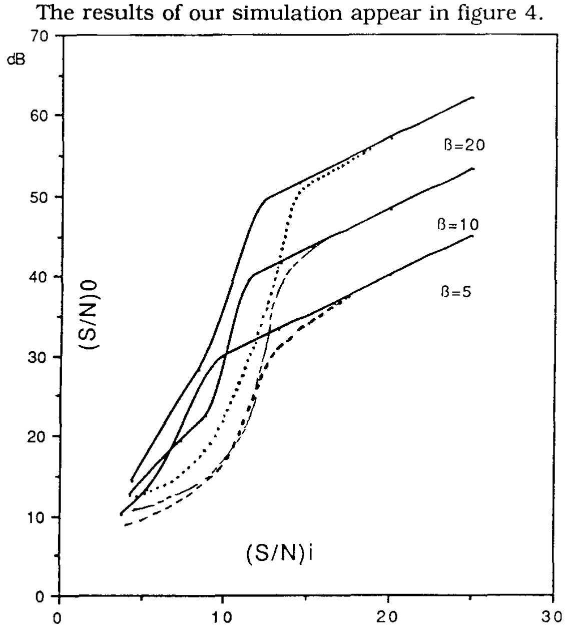

Figure 4: Comparative performance results obtained by simulation of Hagiwara demodulator (dashed lines) and the proposed improved version without threshold extender.

As can be seen from figure 4, our assumption was right. It is preferable to put the range extender before the time differentiator.

3.3 Threshold extension

It is well known that a limiter-discriminator is not optimum near the threshold region $\left[(S/N)_i \approx 10\text{dB}\right]$. That comes from the fact that the limiter discards the information contained in the amplitude fluctuation of the received noisy FM signal. That information could have been used to reduce noise clicks which appear at the output of the demodulator when the $(S/N)_i$ falls below threshold.

Both balanced quadricorrelator and the arctangent type demodulator discussed earlier also discard the information contained in the amplitude fluctuation of the received noisy RF signal. Improved performances could be realized if one takes into account that information. One possible way to realized that improvement would be by the inclusion of an additional circuit like the baseband version of the Frequency Locked Loop (FLL) proposed some years ago by Clarke and Hess [17]. Figure 5 shows such a circuit applied to the improved arctangent type demodulator discussed earlier.

It is easy to verify that the following transfer function hold for this processor

$$V_o(S) = \left[\frac{1/A}{1+\dfrac{S}{A\,E(S)}}\right] S\Phi(S) \qquad (11)$$

where S is the Laplace operator and $V_o(S)$, E(S) and $\Phi(S)$ are respectively the Laplace transforms of $V_o(t)$, e(t) and $\phi(t)$. Improvement results because the output signal is now filtered by a low-pass filter whose cut-off frequency is proportional to the instantaneous envelope e(t) of the received PM/FM noisy signal. As, near threshold, most clicks are accompanied by a sudden decrease of e(t), the power level of those filtered clicks is reduced by the action of the adaptive filter. A threshold extension of a few dB is possible with that extender [17].

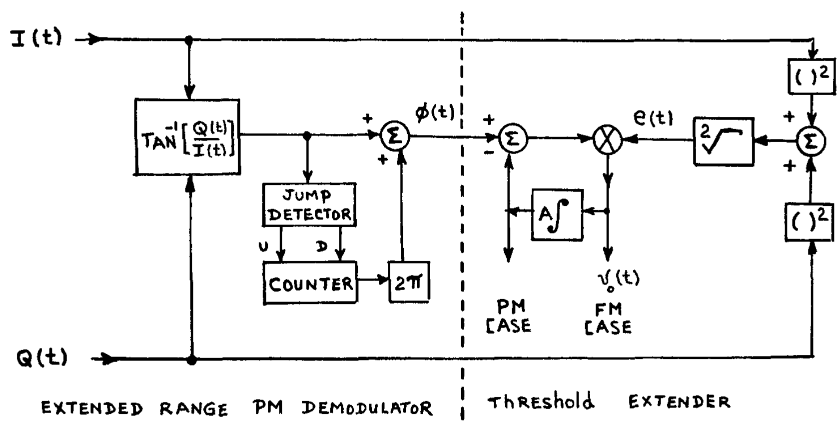

Figure 5: Final version of the extended range arctangent type PM/FM demodulator equipped with a threshold extender.

3.4 Practical considerations

At the beginning of section 3, we assume a perfect local oscillator synchronization ($\Delta\phi(t) = 0$ on figure 1). If it was not the case, it is easy to verify that the signals present at the outputs of the arctangent unwrapper shown in figure 2 are given by

$$\text{PM case:} \quad K_p m(t) + \Delta\phi(t)$$

$$\text{FM case:} \quad K_\omega m(t) + \frac{d}{dt}\Delta\phi(t) \qquad (12)$$

In as much as the offset error $\Delta\phi(t)$ is small and the low-frequency content of m(t) do not interfere with the one of $\Delta\phi(t)$, it is possible to separate $\Delta\phi(t)$ and $\Delta'\phi(t)$ from m(t) and use them as feedback signals to lock the local VCO to the exact required phase and frequency. The same conclusion applies to the final version presented in figure 5 except that, as the "FM output" now precedes the "PM output", no static phase offset $\Delta\phi$ correction can be done from the "PM output". If such a correction is required, one should extract the error signal from the input of the threshold extender, that is at the point labelled $\phi(t)$ in figure 5.

Figure 1 suggests that the quadrature projections I(t) and Q(t) of the complex baseband signal c(t) can be extracted from the PM/FM signal with an analog QAM demodulator. This is not mandatory and even not recommended. Indeed, it is well known that such a direct-conversion or zero IF receiver suffers from imperfections (quadrature error, dc offsets, ...) which cause its performance to depart from ideal. The trend of development in modern communications systems is towards the application of digital processors, which give much higher precision and stability. Complex sampling techniques can be used to supersede analog QAM demodulators. Several approaches have been presented. The interested readers should consult [18-22].

4. Conclusion

It was shown that in PM and FM communication systems a real baseband message is transformed into a complex baseband signal which is a rolled up version of that message on a unity radius cylinder. PM and FM signals could therefore be seen as QAM signals and demodulated as such with a QAM demodulator. Various methods have been presented which permit to unroll the complex received signal in order to give back the original real baseband message. Among them, a modified version of the "ranged extended tangent type FM demodulator" was studied in more detail and found to give excellent results. A special attention was given to digital signal processing.

References

[1] Hagiwara, M. and Nakagawa, M., "Range Extended Tangent Type FM Demodulator with Wide Linear Range Using Digital Signal Processing", Proc. of ISCAS'85, pp. 1481-1482

[2] Hagiwara, M. and Nakagawa, M., "Digital Signal Processing Type Stereo FM Receiver", IEEE Trans. on Consumer Electronics, vol. CE-32, #1, (Feb. 1986), pp. 37-43.

[3] Bedrosian, E., "The Analytic Signal Representation of Modulated Waveforms", Proc. IRE, (Oct. 1962), pp. 2071-2076.

[4] Boutin, N., "Complex Signals: Their processing and Transmission", Proc. of RF Expo East 88, Philadelphia, (Oct. 25-27, 1988), pp. 49-60.

[5] Boutin, N., "Complex Signals: Part IV", RF Design, (May 1990), pp. 65-75.

[6] Park, J.H., "An FM Detector for Low S/N", IEEE Trans. on Communication Technology, vol. Com-18, #2, (April 1970), pp. 110-118.

[7] Denenberg, J.N., "The Power Mean Frequency Estimator: Another Approach to the FM Detector", IEEE Trans. of BTR, vol. BTR-20, #3, (August 1974), pp. 201-204.

[8] "Sine Cosine Frequency Tracker", British patent # 1363396, (Feb. 1973).

[9] "Demodulator for FM Signals", British patent # 1530602, (Oct. 1975).

[10] Vance, I.A.W., "An Integrated Circuit V.H.F. Radio Receiver", The Radio and Electronic Eng., vol. 50, #4, (April 1980), pp. 158-164.

[11] "Transmitter/Receiver for Single Channel Duplex Communication Systems", U.S. patent # 4 238 850, (December 9), 1980.

[12] Gardner, F.M., "Properties of Frequency Difference Detectors", IEEE Trans. on Communication, vol. COM-33, #2, (Feb. 1985), pp. 131-138.

[13] Schaeffer, C.F., "The Zero-Beat Method of Frequency Discrimination", Proc. IRE, vol. 30, (August 1942), pp. 365-367.

[14] Richman, D., "Color Carrier Reference Phase Synchronization Accuracy in NTSC Color Television", Proc. IRE, vol. 42, (Jan. 1954), pp. 106-133.

[15] Brown, R.A. et al, "Some Features of Signal Demodulation Resulting from the Practical Implementation of a Direct Conversion Radio Receiver", Philips Journal of Research, vol. 41, #3, (1986), pp. 219-231.

[16] Thiel, T.E. and Saulnier, G.J., "Simplified Complex Digital Sampling Demodulator", Electronics Letters, vol. 26, #7, (March 1990), pp. 419-421.

[17] Clarke, K.K. and Hess, D.T., "Frequency Locked Loop FM Demodulator", IEEE Trans. on Communication Technology, vol. COM-15, #4, (August 1967), pp. 518-524.

[18] Thiel, T.E. and Saulnier, G.J., "Simplified Complex Digital Sampling", Electronics Letters, vol. 26, #7, (March 1990), pp. 419-421.

[19] Liu, H. et al, "A New Quadrature Sampling and Processing Approach", IEEE Trans. on Aerospace and Electronic Systems, vol. 25, #5, (Sept. 1989), pp. 733-747.

[20] Mitchel, R.L., "Creating Complex Signal Samples from a Band-Limited Real Signal", IEEE Trans. on Aerospace and Electronic Systems, vol. 25, #3, (May 1989), pp. 425-427.

[21] Webb, R.C., "Alternatives to Zero-IF Demodulation", RF Technology Expo'89, (Feb. 14-16, 1989), California, pp. 275-281.

[22] Jackson, M.C., and Matthewson, P., "Digital Processing of Bandpass Signals", GEC Journal of Research, vol. 4, #1, (1986), pp. 32-41.