Wireless Pers Commun
DOI 10.1007/s11277-013-1392-5

A Multiplicative Cancellation Approach to Multipath Suppression in FM Radio

Thomas J. Moir · Archibald M. Pettigrew

Abstract The deceptively simple problem of a single inverted reflection in ordinary frequency modulated (FM) radio is considered. It will be shown that this problem has been overlooked in the literature and causes major breakdown in reception. The problem is known as suppressed-carrier AM-FM (SCAM-FM) and is totally destructive to the received signal. We examine the theory and practical measurements of SCAM and show a solution for reducing its effect.

Keywords FM · Frequency-modulation · Multipath interference · Amplitude-locked loop · Phase locked loop

1 Introduction

There are several breakdown mechanisms in ordinary FM mobile radio. The literature explains such effects as threshold (where the carrier to noise ratio drops bellow an acceptable level), co-channel interference (where more than one transmitter is using the same carrier frequency) and multi-path interference (where a modulated signal is reflected off a nearby building or hill). In most cities, the transmitter power is more than adequate to cope with any problems with threshold and co-channel interference is rarer still since commercial broadcaster are at least aware of this problem and space the transmitters and carrier frequencies accordingly. This leaves the most common problem which is multi-path reflections.

Multipath interference in FM was first analysed by Corrington [1], and is most often found in mobile environments as shown in Fig. 1. The combination of the original FM signal plus a major (and probable many minor) reflections off buildings cause distortion to the demodulated baseband signal which is seen as large spikes [1].

T. J. Moir (✉)
School of Engineering, AUT University, St Pauls Street, Auckland, New Zealand
e-mail: Tom.Moir@aut.ac.nz

A. M. Pettigrew
Ampsys Electronics Ltd., 69 Haco Street, Largs, Scotland

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

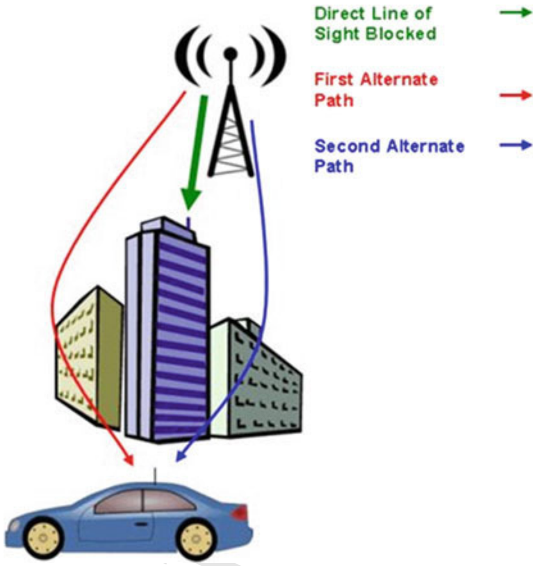


Fig. 1 Multipath interference in a car environment

Since the Corrington paper was published there have been several attempts at reducing the effects of multipath interference. Perhaps one of the most commonest approaches in the literature is to use adaptive equalization [2,3]. Such approaches have proven successful in cases where the reflection is not too severe. Similarly, adaptive filtering methods can be used and combined with so-called “diversity” receivers, whereby improvement is sought with the complexity of using several antennas and receivers [4,5]. One of the earliest approaches was to use the constant modulus algorithm (CMA), but this was found not to respond well to rapid fading [6].

The approach used here also relies on the fact that for ideal FM transmission, the modulus of the waveform will be a constant. However, we consider a special case of multi-path that has been overlooked in the current theory, one single ‘mirror’ reflection with the same amplitude but inverted phase. During short periods of time the carrier disappears completely for this particular case (infinite fading). Maintaining lock with a phase-locked-loop is particularly difficult but is accomplished with the help of an amplitude-Locked-Loop [7]. This has been illustrated elsewhere for synchronous demodulation of double-sideband-suppressed carrier signals [8]. In reference [8] it is shown that a phase-locked loop cannot remain locked during periods of sustained absence of the carrier (that coincides with the free-running frequency of the loop), This is because there is no power at the carrier frequency and hence no signal for the loop to lock into. The inclusion of an amplitude-locked loop enables this problem to be overcome long enough to sustain lock.

The contributions in this paper is first to define mathematically the type of interference that occurs in multipath problems and then to devise a special signal processing (analogue or digital) method which generates an in-phase signal which goes to zero when the interference spike reaches its peak. This is known as a correction signal. Then by the process of multiplication of the correction signal and the interference spikes caused by the process of multipath, a significant reduction (if not total elimination in some cases) in these spikes is obtained. This is made possible because information on the envelope of the composite FM signal plus interference is made use of in generation of the correction signal which cancels the interference. A number of other refinements are also used including the use of the amplitude-locked loop instead of a hard limiter. In fact the amplitude-locked loop is a crucial part of the circuit design.

2 The Basic Problem

Ordinary FM is defined as

$$y(t) = \cos(\omega_c t + \beta \sin(\omega_m t)) \tag{1}$$

where ω_c is the carrier frequency, $\beta = \omega_{pm}/\omega_m$ is the FM modulation index, ω_m is the modulating frequency and ω_{pm} is the peak modulation depth.

Now consider *a single reflection of amplitude m* (normally unity but we keep it general for the time being) delayed by delay τ seconds and added to (1) above giving:

$$f(t) = y(t) + my(t-\tau) \tag{2}$$

$$f(t) = \cos(\omega_c t + \beta \sin(\omega_m t)) + m\cos(\omega_c(t-\tau) + \beta \sin(\omega_m(t-\tau))) \tag{3}$$

Now if this delay is assumed to be $\tau = \pi/\omega_c$,then without losing generality, the delay also applies when substituting π for $(2n-1)\pi, n = 1, 2, 3\ldots$ where n is defined as the *wave number*. This is because in our analysis $\sin((2n-1)\pi) = 0, n = 1, 2, 3\ldots$ and $\cos((2n-1)\pi) = -1, n = 1, 2, 3\ldots$ Then we can expand (3) above in two parts

$$\cos(\omega_c t + \beta \sin(\omega_m t)) = \cos(\omega_c t)\cos(\beta \sin(\omega_m t)) - \sin(\omega_c t)\sin(\beta \sin(\omega_m t)) \tag{4}$$

and

$$\begin{aligned} m\cos(\omega_c(t-\tau) + \beta \sin(\omega_m(t-\tau))) &= m\cos(\omega_c t - \pi)\cos\left(\beta \sin\left(\omega_m t - \frac{\omega_m}{\omega_c}\pi\right)\right) \\ &\quad - m\sin(\omega_c t - \pi)\sin\left(\beta \sin\left(\omega_m t - \frac{\omega_m}{\omega_c}\pi\right)\right) \end{aligned} \tag{5}$$

Now of course $\cos(\omega_c t - \pi) = -\cos(\omega_c t)$ and $\sin(\omega_c t - \pi) = -\sin(\omega_c t)$

To simplify $\cos(\beta \sin(\omega_m t - \frac{\omega_m}{\omega_c}\pi))$ first simply $\sin(\omega_m t - \frac{\omega_m}{\omega_c}\pi)$ and note that $\cos(\frac{\omega_m}{\omega_c}\pi) \approx 1$ since the carrier frequency will be in the region of 88–108 MHz (for broadcast quality FM) and the baseband frequency no larger than 15 kHz. For the worst case scenario this makes $\frac{\omega_m}{\omega_c}\pi$ around 535×10^{-6} radians so that

$$\sin\left(\omega_m t - \frac{\omega_m}{\omega_c}\pi\right) = \sin(\omega_m t)\cos\left(\frac{\omega_m}{\omega_c}\pi\right) - \cos(\omega_m t)\sin\left(\frac{\omega_m}{\omega_c}\pi\right)$$

which is approximately

$$\sin\left(\omega_m t - \frac{\omega_m}{\omega_c}\pi\right) \approx \sin(\omega_m t) - \frac{\omega_m}{\omega_c}\pi\cos(\omega_m t)$$

Hence

$$\begin{aligned} \cos\left(\beta \sin\left(\omega_m t - \frac{\omega_m}{\omega_c}\pi\right)\right) &\approx \cos\left[\beta \sin(\omega_m t) - \beta\pi\frac{\omega_m}{\omega_c}\cos(\omega_m t)\right] \\ &= \cos(\beta \sin(\omega_m t))\cos\left(\beta\pi\frac{\omega_m}{\omega_c}\cos(\omega_m t)\right) + \sin(\beta \sin(\omega_m t))\sin\left(\beta\pi\frac{\omega_m}{\omega_c}\cos(\omega_m t)\right) \end{aligned}$$

which is approximately equal to

$$\approx \cos(\beta \sin(\omega_m t)) + \sin(\beta \sin(\omega_m t))\beta\pi \frac{\omega_m}{\omega_c} \cos(\omega_m t)$$

Similarly we have

$$\sin\left(\beta \sin\left(\omega_m t - \frac{\omega_m}{\omega_c}\pi\right)\right) \approx \sin(\beta \sin(\omega_m t)) - \beta\pi \frac{\omega_m}{\omega_c} \cos(\beta \sin(\omega_m t)) \cos(\omega_m t)$$

adding all the terms in (1) gives

$$\begin{aligned} f(t) \approx{} & \cos(\omega_c t) \cos(\beta \sin(\omega_m t)) - \sin(\omega_c t) \sin(\beta \sin(\omega_m t)) \\ & + m \left\{ -\cos(\omega_c t) \left[\cos(\beta \sin(\omega_m t)) + \beta\pi \frac{\omega_m}{\omega_c} \sin(\beta \sin(\omega_m t)) \cos(\omega_m t) \right] \right. \\ & \left. + \sin(\omega_c t) \left[\sin(\beta \sin(\omega_m t)) - \beta\pi \frac{\omega_m}{\omega_c} \cos(\beta \sin(\omega_m t)) \cos(\omega_m t) \right] \right\} \end{aligned} \tag{6}$$

For the special case of a perfect mirror reflection when m = 1 the above simplifies to

$$f(t) \approx \beta\pi \frac{\omega_m}{\omega_c} \cos(\omega_m t)[-\cos(\omega_c t) \sin(\beta \sin(\omega_m t)) - \sin(\omega_c t) \cos(\beta \sin(\omega_m t))] \tag{7}$$

Further simplifying

$$f(t) \approx -\beta\pi \frac{\omega_m}{\omega_c} \cos(\omega_m t)[\sin(\omega_c t) \cos(\beta \sin(\omega_m t)) + \cos(\omega_c t) \sin(\beta \sin(\omega_m t))]$$

Finally we have

$$f(t) \approx -\beta\pi \frac{\omega_m}{\omega_c} \cos(\omega_m t)[\sin(\omega_c t + \beta \sin(\omega_m t))] \tag{8}$$

The above equation clearly also applies when substituting π for $(2n - 1)\pi$, n = 1, 2, 3...

This is an FM signal which is multiplied by an amplitude term $-\beta(2n-1)\pi \frac{\omega_m}{\omega_c} \cos(\omega_m t)$, n = 1, 2, 3... giving rise to the acronym suppressed-carrier AM-FM or SCAM-FM for short. See the simulation below. This waveform has both FM and AM and was simulated from equation (2) with m = n=1 and not from the simplified model (8) though it has the same form.

We may further examine the effects of this problem in the frequency domain by examining the frequency domain properties of the propagation channel. Although Fig. 2 looks like

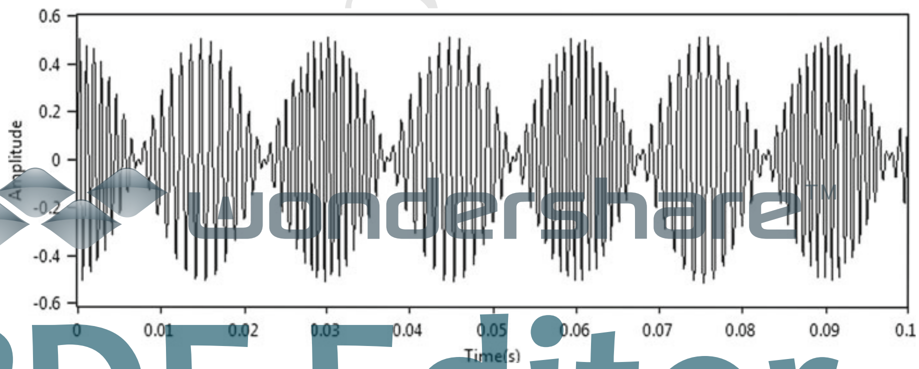


Fig. 2 A simulation of SCAM-FM

Journal: **11277-WIRE** Article No.: **1392** TYPESET DISK LE CP Disp.:**2013/8/29** Pages: **21** Layout: **Small**

normal double-sideband suppressed carrier (DSSC), it should be pointed out that the above waveform is also frequency modulated.

Taking Fourier transforms of (2) gives

$$f(j\omega) = y(j\omega)[1 + me^{-j\omega\tau}] \tag{9}$$

Defining $G(j\omega) = 1 + me^{-j\omega\tau}$ we see that the original FM signal has become convolved (i.e. filtered) by the path transfer function $G(j\omega)$

$$f(j\omega) = G(j\omega)y(j\omega) \tag{10}$$

and we need only find the characteristics of $G(j\omega)$ in order to find the frequency domain properties of SCAM-FM. In fact $G(j\omega)$ has a well known form when m=1 and is easily manipulated.

$$\begin{aligned} G(j\omega) &= 2e^{-j\omega\tau/2}\left[e^{j\omega\tau/2} + e^{-j\omega\tau/2}\right]\Big/2 \\ &= 2e^{-j\omega\tau/2}\cos(\omega\tau/2) \end{aligned}$$

Now substitute $\tau = (2n-1)\pi/\omega_c$, n = 1, 2, 3... and we get

$$G(j\omega) = 2e^{-j\frac{(2n-1)\pi}{2}(\omega/\omega_c)}\cos\left(\frac{(2n-1)\pi}{2}(\omega/\omega_c)\right), \quad n = 1, 2, 3\ldots$$

The magnitude of this becomes

$$|G(j\omega)| = 2\left|\cos\left(\frac{(2n-1)\pi}{2}(\omega/\omega_c)\right)\right|, \quad \mathrm{n} = 1, 2, 3\ldots \tag{11}$$

Which has an absolute valued cosine characteristic of amplitude 2 and it goes to zero at $\omega = \omega_c$. It has a gain of 0.707 (−3 dB down) at $\omega = \omega_c/2$ ie half the carrier frequency. In general zeros of the frequency response occur when

$$\frac{(2n-1)\pi}{2}(\omega/\omega_c) = \frac{(2n-1)\pi}{2}, \quad n = 1, 2, 3\ldots$$

From which $\omega = \omega_c$ Which gives an infinite null at the carrier frequency $\omega = \omega_c$. The frequency response in Hz is illustrated in Fig. 3 below and is independent of the wave number n.

We note from Fig. 3 that the dB gain is minus infinity, totally suppressing the carrier frequency and giving rise to the typical SCAM time-domain composite carrier waveform of Fig. 2.

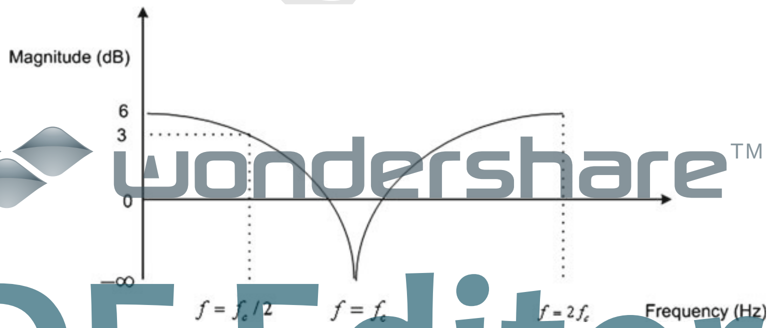


Fig. 3 Magnitude frequency response of SCAM channel

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

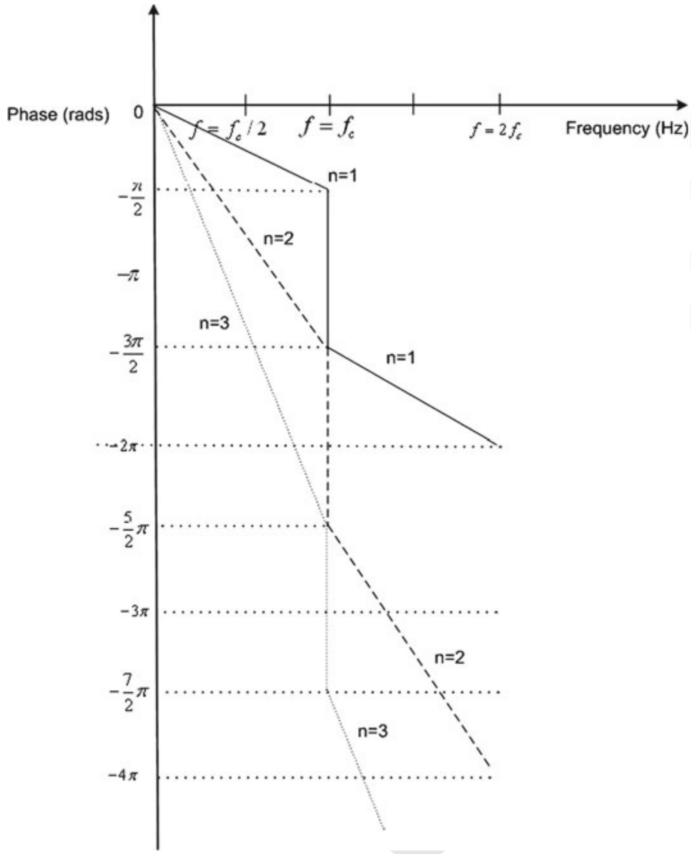


Fig. 4 Phase response of SCAM channel for n = 1,2,3

The phase response is dependant on the wave number n = 1,2,3… and is shown in Fig. 4. Note the rapid phase change of $-\pi$ at the notch frequency f_c for any value of n.

2.1 The Length of the SCAM Zone

Having established that a single reflection at 180° converts the FM carrier into the suppressed carrier amplitude modulated function which is also frequency modulated, it is instructive to develop a more general solution for these conditions.

In the interests of clarity, the value of peak deviation will be used as the modulation depth and defined as $\omega_{pm} = 100\,\text{kHz}$. The carrier frequency will be defined as $\omega_c = 100\,\text{MHz}$. From (8)

$$\begin{aligned} f(t) &= -\beta\pi(\omega_m/\omega_c)(\cos\omega_m t)\sin(\omega_c t + \beta\sin\omega_m t) \\ &= -\pi(\omega_{pm}/\omega_c)(\cos\omega_m t)\sin(\omega_c t + \beta\sin\omega_m t) \\ &= -\pi\,1/1000(\cos\omega_m t)\sin(\omega_c t + \beta\sin\omega_m t) \\ &= -0.003142(\cos\omega_m t)\sin(\omega_c t + \beta\sin\omega_m t) \end{aligned}$$

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

So for a carrier received at a signal strength of 1 millivolt, after the attenuation the carrier strength will be reduced to 3.142 microvolt. If we now apply the same argument for 3π

$$\mathrm{f(t)} = (3)(-0.003142(\cos\omega_m t)\sin(\omega_c t + \beta\sin\omega_m t))$$

$$\mathrm{f(t)} = 0.009426(\cos\omega_m t)\sin(\omega_c t + \beta\sin\omega_m t))$$

So the signal strength increases by (2n−1) as the delay distance increases.where n is the wave number.

If the delay length is not exactly 180°, then there will still be a certain frequency offset from the carrier frequency where the same phase discontinuity will occur. This phase discontinuity will occur higher and higher up the modulating cycle either in the positive or the negative direction. When the distance is such that no discontinuity appears in the carrier then this is the end of the SCAM zone and normal amplitude modulation returns. This point is reached when the peak amplitude modulated signal reaches twice the value of the attenuated carrier with pure symmetric SCAM.

In the case of a delay of 180° (1.5 m) this evaluates to 1/1,000 of the half wavelength or 1.5 mm. The peak-to-peak value is twice this value or 3 mm.

$$\lambda_{SCAM} = \lambda_{wavelength}\omega_{pm}/\omega_c$$

Or more generally

$$\lambda_{SCAM}(n) = (2n-1)\lambda_{wavelength}\omega_{pm}/\omega_c \quad \mathrm{n} = 1, 2, 3\ldots \tag{12}$$

There is a limit in the length of the SCAM zone that has been found experimentally

$$\lambda_{SCAM}(\max) = \lambda_{wavelength}1/1.57 \tag{13}$$

So, with a wavelength of 3 m the maximum length of SCAM zone can be 3/1.57 m or 191 cm. at the delay distance of 751.5 m.

The analysis shows that there is a zone of catastrophic failure in FM reception called the SCAM zone where the carrier is infinitely attenuated by a short distance reflection and that this distance is a function of the ratio of the peak modulation to the carrier frequency and is proportional to the delay distance up to about 750 m where the zone is 191 cm in length.

The distortion caused by the phase discontinuity cause severe distortion that will be analysed in the next section.

3 Demodulation in the SCAM Zone

There are many issue to consider when examining how an FM demodulator will treat the basic SCAM equation (8). For instant when a limiter is present prior to say a phase-locked-loop and there is additive noise, the period of time when the carrier goes to zero will be chaotic to the demodulation process. This is because a limiter will only amplify the noise when no carrier is present. To simplify the procedure we consider the absence of noise and that initially the phase-locked loop is ideal (i.e. it cannot respond to amplitude variations). Equation (8) cannot be used directly in following analysis since when m=1 the problem is essentially 'singular'. Instead we must consider when m is just less than unity. Therefore returning to (30) with $\tau = \pi/\omega_c$ we get

$$\begin{aligned} f(t) = \cos(\omega_c t + \beta\sin(\omega_m t)) + m\cos[\omega_c t + \beta\sin(\omega_m t) \\ + \beta\sin(\omega_m(t - \pi/\omega_c)) - \pi - \beta\sin(\omega_m t)] \end{aligned} \tag{14}$$

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

This we can simplify using trig identities to give

$$f(t) = [1 + m\cos(\omega_d(t))]\cos(\omega_c t + \beta\sin(\omega_m t)) - m\sin(\omega_d(t))\sin(\omega_c t + \beta\sin(\omega_m t)) \tag{15}$$

where we define

$$\psi(t) = \beta\sin(\omega_m(t - \pi/\omega_c)) - \beta\sin(\omega_m t) - \pi \tag{16}$$

The Eq. (16) $\psi(t)$ can be simplified (see Appendix 1) and will be shown to be related to twice the modulating frequency. Looking at Eq. (15) for the composite carrier we can simplify this by using the well known trig identity which follows from a right-angled triangle. We can express (15) in the simple form

$$a\cos(x) + b\sin(x) = c\cos(x - \phi)$$

where

$$\phi = \tan^{-1}[b/a], \quad c = \sqrt{a^2 + b^2}$$

with

$$a = [1 + m\cos(\psi(t))], \quad b = -m\sin(\psi(t))$$

Hence the composite waveform (15) can be written as

$$f(t) = r(t)\cos[\omega_c t + \beta\sin(\omega_m t) - \phi(t)] \tag{17}$$

The amplitude part $r(t)$ equal to

$$r(t) = \sqrt{1 + 2m\cos(\psi(t)) + m^2} \tag{18}$$

and the phase

$$\phi(t) = \tan^{-1}\left[-\frac{m\sin(\psi(t))}{1 + m\cos(\psi(t))}\right] \tag{19}$$

3.1 Phase-Locked-Loop Demodulated Output

The (ideal) phase-locked loop output p(t) is given by the derivative of the phase term of the composite waveform $f(t)$ in (17) i.e.

$$p(t) = \frac{d}{dt}[\omega_c t + \beta\sin(\omega_m t) - \phi(t)]$$

(The ideal phase-locked loop will not respond to the amplitude variations but is helped by the fact that the input composite FM signal has a constant amplitude obtained by using an amplitude-locked-loop (ALL) and not a hard-limiter)

The first of these derivatives becomes ω_c which we shall ignore as this is just a dc component. The second of these derivatives is

$$\frac{d}{dt}[\beta\sin(\omega_m t)] = \beta\omega_m\cos(\omega_m t)$$

and this is the ideal PLL output when there is no SCAM interference.

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

Since $\tan^{-1}(-x) = -\tan^{-1}(x)$, the final derivative is given by

$$-\frac{d}{dt}[\phi(t)] = -\frac{d}{dt}\tan^{-1}\left[-\frac{m\sin(\psi(t))}{1+m\cos(\psi(t))}\right]$$
$$= \frac{d}{dt}\tan^{-1}\left[\frac{m\sin(\psi(t))}{1+m\cos(\psi(t))}\right]$$

and after some algebra (Appendix 2) this leads to the solution

$$-\frac{d}{dt}[\phi(t)] = \left[\frac{m^2+m\cos(\psi(t))}{1+2m\cos(\psi(t))+m^2}\right]\beta\omega_m[\cos(\omega_m(t-\pi/\omega_c)) - \cos(\omega_m t)]$$

Adding all three terms reveals the solution for one reflection for a delay of π/ω_c radians to be:

$$\begin{aligned} p(t) &= \frac{(1+m\cos(\psi(t))}{(1+2m\cos(\psi(t))+m^2)}\beta\omega_m\cos(\omega_m t) \\ &+ \frac{(m^2+m\cos(\psi(t))}{(1+2m\cos(\psi(t))+m^2)}\beta\omega_m\cos(\omega_m(t-\pi/\omega_c)) \end{aligned} \tag{20}$$

3.2 Simplification of the Above Phase-Locked Loop Expression

To simplify (20) above we need to evaluate $\cos(\omega_m(t-\pi/\omega_c))$ on the right-hand side of the equation.

Expanding gives

$$\cos(\omega_m(t-\pi/\omega_c)) = \cos(\omega_m t)\cos(\pi\omega_m/\omega_c) + \sin(\pi\omega_m/\omega_c)\sin(\omega_m t)$$

(since $\pi\omega_m/\omega_c$ is a small angle we can approximate the above by using $\cos(x) \approx 1 - \frac{x^2}{2}$ and $\sin(x) \approx x - \frac{x^3}{6}$ but we shall ignore the cubed term)

Hence by defining $k = \frac{\pi\omega_m}{\omega_c}$ we have

$$\cos(\omega_m(t-\pi/\omega_c)) = \cos(\omega_m t) - \frac{k^2}{2}\cos(\omega_m t) + k\sin(\omega_m t)$$

Similarly we can simplify the term $\cos(\omega_d(t))$ by using the expression in Appendix 1.

$$\cos(\psi(t)) \approx -1 + \frac{1}{4}(\beta k)^2[1+\cos(2\omega_m t)]$$

Now define the denominator term from the phase-locked loop output p(t) as

$$d(t) = 1 + 2m\cos(\psi(t)) + m^2$$

and substituting $\cos(\omega_d(t))$ gives

$$d(t) \approx (1-m)^2 + \frac{m}{2}(\beta k)^2[1+\cos(2\omega_m t)]$$

and define the numerator term from p(t) in (20) as

$$n(t) = m^2 + m\cos(\psi(t))$$

which becomes

$$n(t) \approx m^2 - m + \frac{m}{4}(\beta k)^2[1+\cos(2\omega_m t)]$$

Using the above results the simplified phase-locked loop output results in

$$p(t) \approx \beta\omega_m \cos(\omega_m(t)) \left[1 - \frac{k^2}{2}\frac{n(t)}{d(t)}\right] + k\frac{n(t)}{d(t)}\beta\omega_m \sin(\omega_m(t))$$

This is a sum of the ideal phase-locked loop output with m=0 and a quadrature term which gives rise to spikes. The term in $\frac{k^2}{2}$ can be ignored under normal conditions leaving the final simplified form:

$$p(t) \approx \beta\omega_m \cos(\omega_m(t)) + k\frac{n(t)}{d(t)}\beta\omega_m \sin(\omega_m(t)) \tag{21}$$

3.3 Summary of Findings

The simplified phase-locked loop output when m is just less than unity (say m=0.999) is given by

$$p(t) \approx \beta\omega_m \cos(\omega_m(t)) + k\frac{n(t)}{d(t)}\beta\omega_m \sin(\omega_m(t))$$

where

$$k = \frac{\pi\omega_m}{\omega_c}$$

$$n(t) \approx m^2 - m + \frac{m}{4}(\beta k)^2[1 + \cos(2\omega_m t)]$$

$$= (1-m)^2 + (m-1) + \frac{m}{4}(\beta k)^2[1 + \cos(2\omega_m t)]$$

and

$$d(t) \approx (1-m)^2 + \frac{m}{2}(\beta k)^2[1 + \cos(2\omega_m t)]$$

This type of phase-locked loop output contains the demodulated baseband with spikes which occur *at the zero-crossings* of the waveform. The spikes are particularly destructive to any audio waveform.

3.4 Height of the Spikes

The height of the spikes follows from the expression of the phase-locked loop output. The second part of the phase-locked loop output expression is:

$$p_2(t) = k\frac{n(t)}{d(t)}\beta\omega_m \sin(\omega_m t)$$

(since the first part contains only a pure cosine wave)

The spikes have equal height and occur at the maxima and minima of the sin function. That is when $\omega_m t = i\frac{\pi}{2}, i = 1, 3, 5 \ldots$ Substituting this angle into n(t) and d(t) gives the magnitude

$$p_2(\max) = \left|k\frac{(1-m)^2 + (m-1)}{(1-m)^2}\beta\omega_m\right|$$

$$p_2(\max) = \frac{\pi\omega_m}{\omega_c}\frac{m}{(1-m)}\omega_{pm}$$

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

3.5 Illustrative Example

The commercial FM frequency band in the UK is from 88–108 MHz. Any multipath composite carrier signal in these range of frequencies is converted down to the common intermediate frequency of 10.7 MHz which is where the phase-locked loop is centred. For simplification the simulation does not consider the down-conversion part of this process. Consider a single reflection problem when the modulating frequency is 4 kHz, the peak modulating frequency is 100 kHz, the IF is 10.7 MHz and m = 0.999. For this problem $\beta = 25(\omega_{pm} = 2\pi \times 100\,\text{kHz})$ and the theoretical phase-locked loop output when plotted is shown in Fig. 5 below.

The spikes are clearly visible at the zero-crossings of the demodulated signal. Note the size of the spike relative to the signal size. The spike height follows from

$$p_2(\max) = \frac{\pi \omega_m}{\omega_c} \frac{m}{(1-m)} \omega_{pm}$$

which becomes

$$p_2(\max) = 7.3717 \times 10^5$$

Of course in a practical circuit there will be finite bandwidth of the phase-locked loop and a saturation limit of the amplifiers. However, any type of impulsive noise will have a wide spectrum across the baseband spectrum and is difficult to remove. In fact it has the effect of pushing up the noise-floor of the demodulated signal and hence reducing signal to noise ratio.

4 Suppression of the Zero-Crossing Spikes

The method used here is quite complex but can be summarised in a simpler way by first noticing that from Eq. (20) for the demodulated output we know that

$$\begin{aligned} d(t) &= 1 + 2m\cos(\psi(t)) + m^2 \\ &= r^2(t) \end{aligned}$$

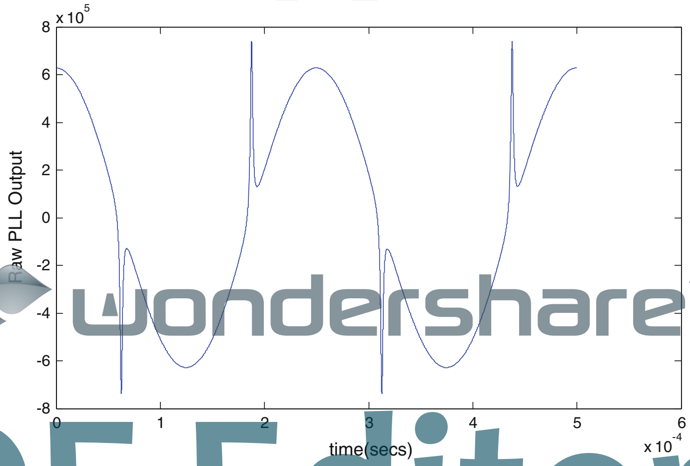


Fig. 5 Phase-locked loop output for a single 180° reflection (SCAM)

Journal: **11277-WIRE** Article No.: **1392** ☐ TYPESET ☐ DISK ☐ LE ☐ CP Disp.:**2013/8/29** Pages: **21** Layout: **Small**

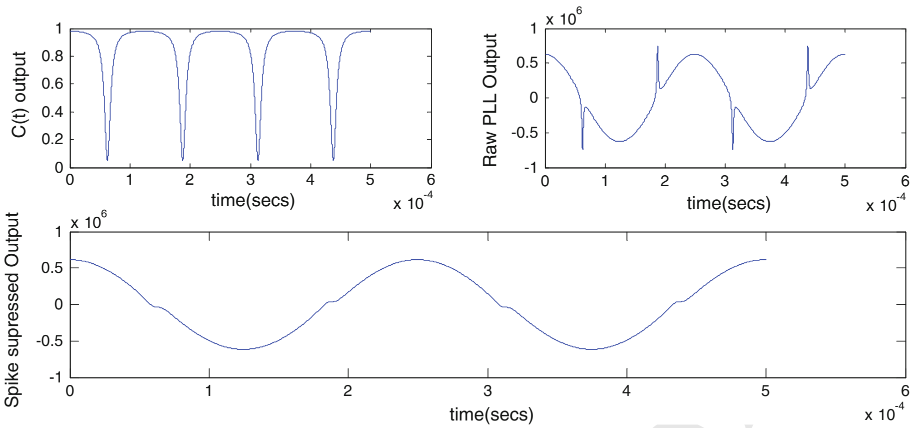


Fig. 6 Shows a good cancellation effect of the spikes

That is, the denominator of the phase-locked loop output is exactly the same as the envelope of the composite multipath FM waveform squared (18). Therefore it is easy to obtain a good estimate of $d(t)$ from the envelope of the composite FM waveform by squaring and lowpass filtering. Once we find this we can construct a special waveform called the "correction" waveform $c(t)$ given by

$$c(t) = \frac{d(t)}{[d(t) + \delta]} \tag{22}$$

where δ is a small positive constant offset voltage. The waveform produces spikes which are time-aligned to the phase-locked loop spikes (which in turn are time-aligned to the nulls in the composite carrier waveform), except that they reach down towards zero. In fact the narrower and closer to zero that these spikes are, the better performance, though there are several trade-offs. When multiplied by the phase-locked loop output i.e. $c(t)p(t)$,the net effect is a much purer waveform with spike suppression.

Using the same example 3.5 above, we can see the effect in Fig. 6 below.

In Fig. 6 we see the correction waveform $c(t)$ the original phase-locked loop output with spikes and the product $c(t)p(t)$.

It can be seen that the spikes have been totally removed except for some small residual caused by the fact that the spikes of c(t) did not reach exactly zero. If the spikes are too wide then there will be dead-zone at the zero-crossings of $c(t)p(t)$. Too narrow and there will not be enough cancellation. The system works by simply multiplying by zero at the correct time.

This is therefore a multiplicative cancellation system and not convolutive as with the case of adaptive filters.

The key to good cancellation is the shape of the pulse and the nearness to zero. For example, $\delta = 0.000018$ in Fig. 6. If the offset value is made too big then the corrective spikes will be too fat as illustrated below in Fig. 7 when $\delta = 0.00018$, ten times larger.

When the corrective spikes are too fat this will result in distortion as shown above.

Likewise, if the offset is made too small then we have the situation as shown in Fig. 8 where not enough cancellation is made. The pulses do not reach zero and there are residual spikes left.

To get a greater insight to the benefits, consider a 2 kHz baseband signal with $\beta = 50$. The demodulated signal in the frequency-domain is shown before and after cancellation (Fig. 9).

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

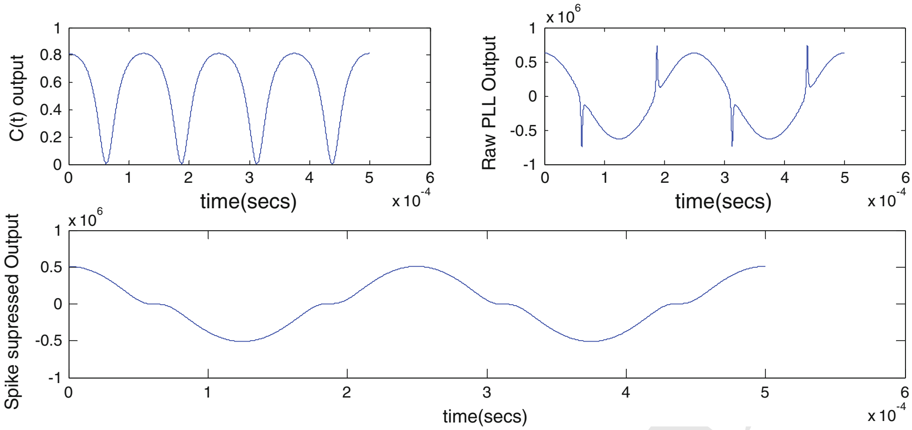


Fig. 7 Shows distortion when offset δ is made too large

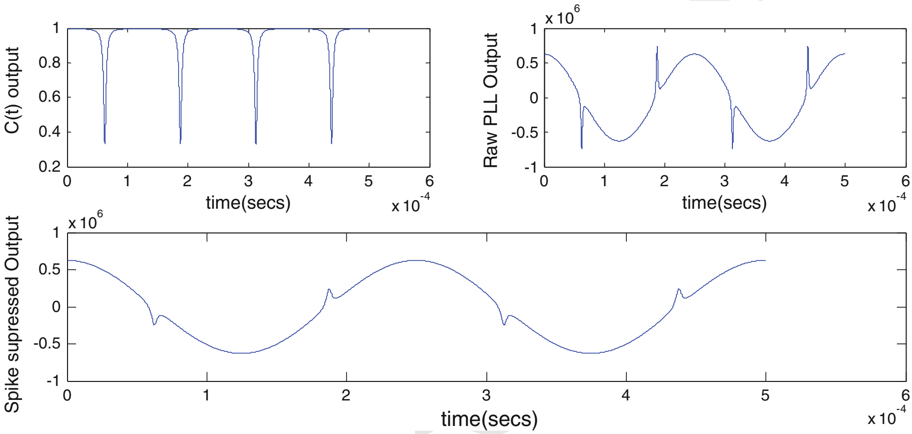


Fig. 8 Shows insufficient cancellation when offset δ is made too small

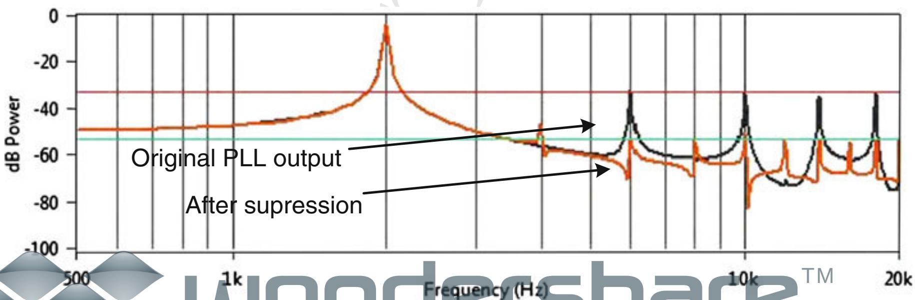


Fig. 9 Spectrum of demodulated phase-locked loop (PLL) signal before and after multipath. suppression

It can be seen that the reduction in harmonic power is by up to as much as 20 dB with detrimental effect on the spectrum of the signal itself. The suppression occurs right across the audio frequency spectrum.

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

5 Analogue System Realization

Although in theory the previous sections describe a mathematical method of removing multipath propagation, practically, the circuit realization is far from easy. Figure 10 shows the block-diagram of the complete system working at the intermediate frequency (IF) of typical high-quality FM.

The first thing which is noticeably different is that there is no hard-limiter in the system at all. This is because the envelope needs to be preserved in order to obtain the envelope which is squared and filtered to obtain $d(t)$. The envelope comes from a slow-acting automatic-gain control circuit (AGC). The AGC output is fed into a high bandwidth amplitude-Locked Loop. The amplitude-locked loop has been discussed elsewhere [7,8] and is a linear circuit which flattens the envelope in a similar fashion to that of a hard-limiter but without amplifying the zero-cross noise. In Fig. 10 the AGC has a wide dynamic range of around 120 dB whereas the amplitude-locked loops dynamic range is around 20 dB. A hard-limiter relies on the zero-crossings of the composite carrier waveform and can have the effect at low signal-to-noise ratios of amplifying any additive noise. Figures 11 and 12 show the block diagrams of an amplitude-locked-loop and a hard-limiter respectively.

The amplitude-locked loop works by maintaining negative feedback and high open-loop gain around an integrator. Since the input is ac, the squarer acts as a transducer which measures the power (or variance) in the signal and adjusts the amplitude of the output to have a carrier power defined by the setpoint. This in turn defines the amplitude of the amplitude-locked loop

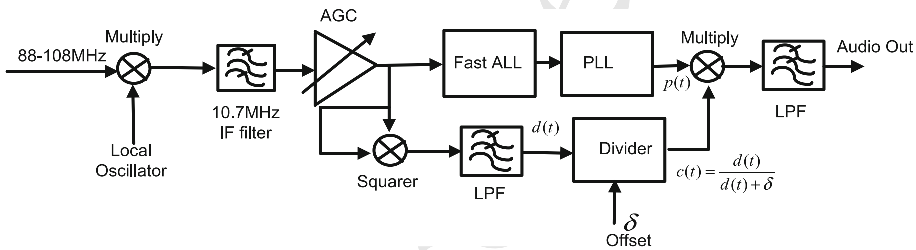
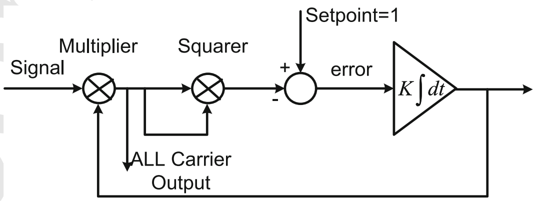


Fig. 10 Block diagram of complete demodulation process

Fig. 11 Amplitude locked-Loop



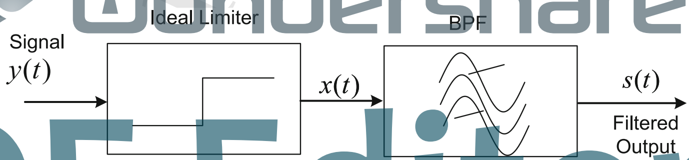


Fig. 12 Hard limiter block diagram

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

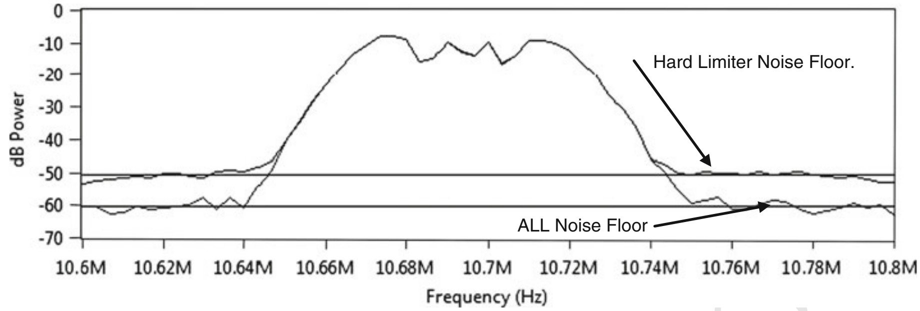


Fig. 13 Comparison of carrier amplitude-locked loop (ALL) and hard limiter noise floor

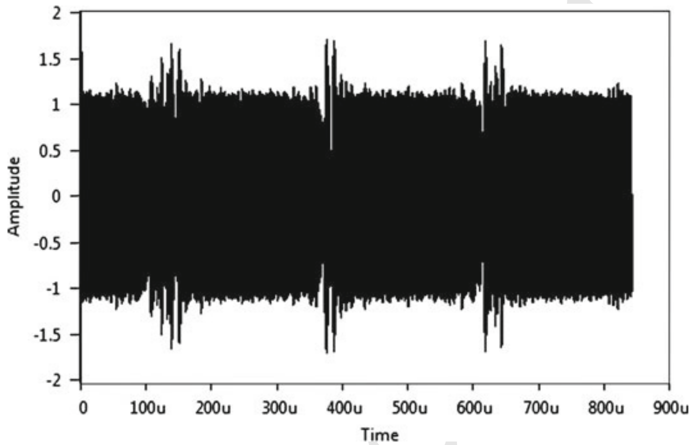


Fig. 14 Amplitude locked-loop recovered carrier

output which consists of a constant amplitude carrier. Of course if there are gaps in the carrier waveform where the carrier goes to zero (as is the case with multipath), then the amplitude-locked loop does not attempt to amplify the noise up to the same level as the envelope (as with an AGC). Instead, it jumps "out of lock" in a similar fashion to that of it's dual circuit, the phase-locked-loop. This is a desirable property unlike the hard-limiter of Fig. 12 which will react to additive noise (or dc-offsets) and produce a higher noise-floor than that of an amplitude-locked loop. This can be shown with a simulation of an ideal amplitude-locked loop and ideal Hard Limiter.

Consider the case of an FM modulated signal of 2 kHz with $\beta = 5$. There is no multipath spikes present but additive channel noise has been added which results in an overall carrier-to-noise ratio of 17 dB. This is before the onset of thresholding. Figure 13 shows a comparison of the IF spectrum at the output of an amplitude-locked loop and hard limiter.

The amplitude-locked loop noise floor is approximately 9 dB down on the equivalent limiter noise floor. Thus it is justifiable in using the circuit in replace of a hard limiter in the overall system.

The amplitude-locked loop recovered carrier output is shown in Fig. 14.

In fact the compression of the carrier increases the power at the carrier (of course a hard limiter also does this) but with the result of a smaller noise-floor.

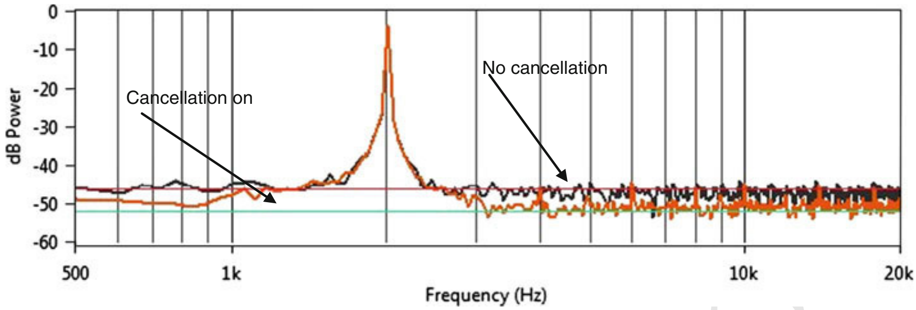


Fig. 15 Effect on baseband noise-floor of cancellation of multipath and noise

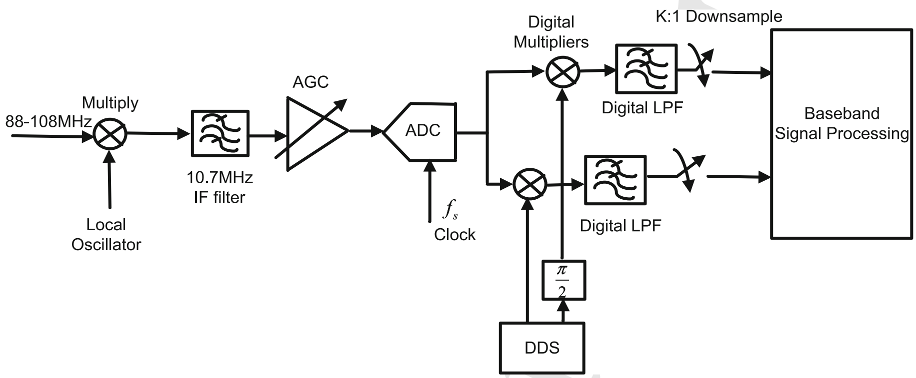


Fig. 16 Software radio architecture

The corresponding demodulated output before and after cancellation of the spikes *when multipath is present* is shown in Fig. 15. The reduction is of course much less than in the ideal noise-free case and shows a reduction of 6 dB in the noise-floor after cancellation of the spikes.

A similar (but not identical) approach was previously shown in [9]. For that paper a different method was used to produce the correction term to cancel the spikes and the method will not work on multipath propagation. A similar yet different approach was used for co-channel interference in Laser Doppler Velocimetry[10]. This is the first time such an approach has been applied to multipath FM (Fig. 16).

6 Software Radio Approach

We consider a software approach to the demodulation process. Such an approach is now commonplace and finding its way into many real FM radios [11].

In this approach the conversion from analogue to digital occurs at IF rather than baseband and the down-conversion to baseband is performed digitally including a sample rate reduction. The above architecture is free of imbalance problems between I and Q. There is an extension of the above architecture whereby the entire FM spectrum is bandlimited (88–108 MHz) by a 20 MHz bandpass filter and then sampled and down-converted in a similar manner to the above. A sampling rate of 80 MHz is often used for this application and this aliases the original

Journal: 11277-WIRE Article No.: 1392 TYPESET DISK LE CP Disp.:2013/8/29 Pages: 21 Layout: Small

FM band which is centred on 98 MHz down to 18 MHz. This is close to quarter sample rate of the converter and gives good separation of the various aliased images. Whichever method is used will result in the same end-result which are two components I and Q. For our SCAM-FM problem we have an analogue composite FM IF signal (with multipath) from the AGC given by:

$$f(t) = r(t)\cos[\omega_c t + \beta \sin(\omega_m t) - \phi(t)] \tag{23}$$

To find the I and Q components we first convert the above signal to a sampled signal given by

$$f(nT_s) = r(nT_s)\cos[\omega_c nT_s + \beta \sin(\omega_m nT_s) - \varphi(nT_s)], \quad \text{n} = 0, 1, 2, 3\ldots \tag{24}$$

and $T_s = \frac{1}{f_s}$ is the sampling interval in seconds. For mathematical convenience we drop the sampling interval write the above as

$$f(n) = r(n)\cos[\omega_c n + \beta \sin(\omega_m n) - \varphi(n)], \quad \text{n} = 0, 1, 2, 3\ldots \tag{25}$$

We now digitally convert to baseband by multiplying $f(n)$ by $\cos(\omega_c n)$ and $\sin(\omega_c n)$. This gives us after filtering out the $(2\omega_c n)$ terms:

$$I(n) = r(n)\cos[\beta \sin(\omega_m n) - \phi(n)] \tag{26a}$$

$$Q(n) = r(n)\sin[\beta \sin(\omega_m n) - \phi(n)] \tag{26b}$$

which is all that is required for demodulation. A similar result would be obtained using any of the other architectures. Down-sampling would normally now occur to lessen the computational load but for mathematical purposes of demodulation this need not be used as the same result will be obtained with or without a sample rate reduction.

The FM demodulated signal which is the digital equivalent of the analogue phase-locked loop output is now given by

$$p(n) = \frac{d}{dn}\tan^{-1}\left(\frac{Q(n)}{I(n)}\right) \tag{27}$$

This is easily found to be

$$p(n) = \left(\frac{I(n)\dot{Q}(n) - \dot{I}(n)Q(n)}{I^2(n) + Q^2(n)}\right) \tag{28}$$

Note that in the above $I^2(n) + Q^2(n) = r^2(n)$ (from the expressions for I and Q)

A number of possible solutions are available to solve the above expressions for $p(n)$ including using FIR filters to produce a band-limited differentiator for instance. Here we consider a simple solution to show the basic idea. Approximate a differentiator as a first difference:

$$\dot{I}(n) = \frac{d}{dn}I(n) \approx I(n) - I(n-1) \tag{29}$$

$$\dot{Q}(n) = \frac{d}{dn}Q(n) \approx Q(n) - Q(n-1) \tag{30}$$

Substituting into (28) gives

$$p(n) = \left(\frac{Q(n)I(n-1) - I(n)Q(n-1)}{I^2(n) + Q^2(n)}\right) \tag{31}$$

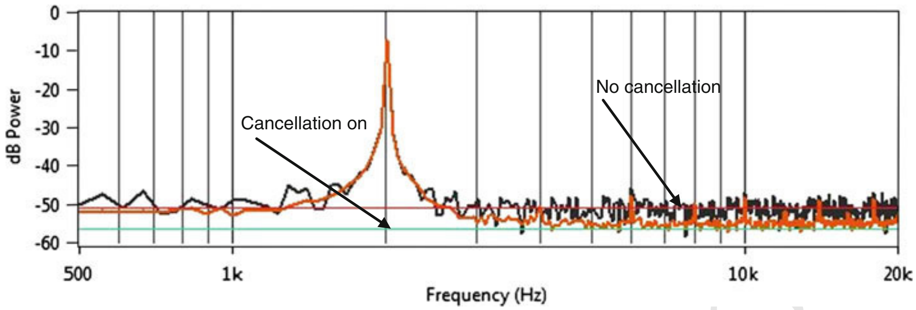


Fig. 17 Software radio: effect on baseband noise-floor of cancellation of multipath and noise

To nullify a spike we need to generate the digital equivalent of a spike going to zero at the right time (the zero-crossing of the modulating frequency). The analogue correction signal is:

$$c(t) = \frac{d(t)}{[d(t) + \delta]}$$

But since $d(t) = r^2(t)$ ie the envelope squared which is the same as the AGC output then clearly the digital equivalent is

$$c(n) = \frac{r^2(n)}{[[r^2(n) + \delta]} \tag{32}$$

and $r^2(n) = I^2(n) + Q^2(n)$ is easily found since I and Q are already known.

$$c(n) = \frac{I^2(n) + Q^2(n)}{[I^2(n) + Q^2(n) + \delta]} \tag{33}$$

The above expression is the digital equivalent of $c(t)$ and no amplitude-locked loop is required.

Now considering the same problem as discussed previously (Fig. 15), but now for the software radio case (an FM modulated signal of 2 kHz with $\beta = 5$. There is multipath present and additive channel noise has been added which results in an overall carrier-to-noise ratio of 17 dB)

It is interesting here to note that the spectrum *before* cancellation in Fig. 17 is equivalent in terms of noise-floor as the previous analogue noise-floor case *after* cancellation. An extra 6 dB reduction in noise-floor is obtained after cancellation using the software approach. This indicates a 12 dB reduction with comparison to an analogue signal with no cancellation. Of course with the digital receiver the sampling process is considered ideal and no quantization noise is modelled in the simulation.

7 Conclusions

A new approach to suppression of multipath noise has been shown which does not involve convolutive or adaptive filtering. Both an analogue and software radio architecture has been presented which shows that the analogue receiver with canceller is superior in performance to one without, and the software approach has the added advantage of a simple solution and even lower noise floor. The method has a wide range of applications mainly in mobile FM radio systems where the sound quality is of great concern.

Appendix 1. Simplification of $\psi(t)$

The difference frequency function is given by (16)

$$\psi(t) = \beta \sin(\omega_m(t - \pi/\omega_c)) - \beta \sin(\omega_m t) - \pi \tag{34}$$

Now the fist sine term in (34) can be written as

$$\beta \sin(\omega_m(t - \pi/\omega_c)) = \beta \left[\sin(\omega_m t) \cos\left(\frac{\pi \omega_m}{\omega_c}\right) - \cos(\omega_m t) \sin\left(\frac{\pi \omega_m}{\omega_c}\right) \right]$$

which approximates for small angles of $k = \frac{\pi \omega_m}{\omega_c}$ to be

$$\beta \sin(\omega_m(t - \pi/\omega_c)) \approx \beta \left[\sin(\omega_m t) \left(1 - \frac{k^2}{2}\right) - k \cos(\omega_m t) \right]$$

Adding all terms gives in (34)

$$\psi(t) \approx \beta \sin(\omega_m t) - \beta k \cos(\omega_m t) - \beta \sin(\omega_m t) - \frac{k^2}{2} \beta \sin(\omega_m t) - \pi$$

or

$$\psi(t) \approx -[\pi + \beta k \cos(\omega_m t)] - \frac{k^2}{2} \beta \sin(\omega_m t)$$

Ignoring the second term above as it will be small we get the result

$$\psi(t) \approx -(\pi + \beta k \cos(\omega_m t))$$

When taking the cosine of $\psi(t)$ we can simplify accordingly

$$\cos \psi(t) \approx \cos[-(\pi + \beta k \cos(\omega_m t))] = -\cos[\beta k \cos(\omega_m t))]$$

and this can be further simplified for small angles to be

$$\begin{aligned} \cos(\psi(t)) &= -\cos[\beta k \cos(\omega_m t))] = -\left[1 - \frac{1}{2}(\beta k)^2 \cos^2(\omega_m t)\right] \\ &= -\left[1 - \frac{1}{4}(\beta k)^2 [1 + \cos(2\omega_m t)\right] \\ &= -1 + \frac{1}{4}(\beta k)^2 [1 + \cos(2\omega_m t)] \end{aligned}$$

Hence this function is related to twice the modulating frequency.

Appendix 2. Differentiation of $\frac{d}{dt} \tan^{-1}\left[\frac{m \sin(\psi(t))}{1+m \cos(\psi(t))}\right]$

Recall that the differentiation of $\frac{d}{dt} \tan^{-1}[f(t)] = \frac{1}{1+[f(t)]^2} f'(t)$ where the shorthand notation is used here $f'(t) = \frac{d}{dt} f(t)$.

Journal: 11277-WIRE Article No.: 1392 ☐ TYPESET ☐ DISK ☐ LE ☐ CP Disp.:2013/8/29 Pages: 21 Layout: Small

To differentiate a function of the form $\frac{u(t)}{v(t)}$ we use the quotient rule

$$\frac{d}{dt}\left(\frac{u(t)}{v(t)}\right) = \frac{v(t)du(t) - u(t)dv(t)}{v^2(t)}$$

so that

$$\begin{aligned}
&= \frac{d}{dt}\tan^{-1}\left[\frac{m\sin(\psi(t))}{1+m\cos(\psi(t))}\right] \\
&= \left[\frac{1}{1+\{\frac{m\sin(\psi(t))}{1+m\cos(\psi(t))}\}^2}\right]\frac{d}{dt}\left[\frac{m\sin(\psi(t))}{1+m\cos(\psi(t))}\right] \\
&= \left[\frac{(1+m\cos(\psi(t))^2}{1+2m\cos(\psi(t))+m^2}\right]\frac{d}{dt}\left[\frac{m\sin(\psi(t))}{1+m\cos(\psi(t))}\right] \\
&= \left[\frac{(1+m\cos(\psi(t))^2}{1+2m\cos(\psi(t))+m^2}\right] \\
&\quad\times\left[\frac{(1+m\cos(\psi(t)))m\cos(\psi(t)) - m\sin(\psi(t))(0-m\sin(\psi(t))}{(1+m\cos(\psi(t)))^2}\right]\frac{d}{dt}\psi(t) \\
&= \left[\frac{m^2+m\cos(\psi(t))}{1+2m\cos(\psi(t))+m^2}\right]\frac{d}{dt}\psi(t)
\end{aligned}$$

Now

$$\psi(t) = \beta\sin(\omega_m(t-\pi/\omega_c)) - \beta\sin(\omega_m t) - \pi$$

so that its derivative is given by

$$\frac{d}{dt}\psi(t) = \beta\omega_m[\cos(\omega_m(t-\pi/\omega_c)) - \cos(\omega_m t)]$$

hence

$$\begin{aligned}
&\frac{d}{dt}\tan^{-1}\left[\frac{m\sin(\psi(t))}{1+m\cos(\psi(t))}\right] \\
&\quad= \left[\frac{m^2+m\cos(\psi(t))}{1+2m\cos(\psi(t))+m^2}\right]\beta\omega_m[\cos(\omega_m(t-\pi/\omega_c))-\cos(\omega_m t)]
\end{aligned}$$

References

1. Corrington, M. S. (1945). Frequency modulation distortion caused by multi-path transmission. *Proceedings of the IRE*, *30*, 878–891.
2. Kammeyer, K. D. (1985). Equalization problems in digital FM receiver. *Signal Processing*, *9*, 263–276.
3. Kammeyer, K. D., Mann, R., & Toberegt, W. (1987). A modified adaptive FIR equalizer for multipath echo cancellation in FM transmission. *IEEE Journal, Selected areas of Communications*, *SAC–5*(2), 226–237.
4. Tingley-Robert, D. (1997). Reducing multipath fading using adaptive filtering. US Patent application No US/08/307300, Filed 16th Sept 1994 issued 9th Dec.
5. Schnepp, H., Muller, T. & Luy, J.-F. (2001). Channel-diversity in software autoradio. In *Proceedings vehicular technology conference*, VTC-Spring, Rhodes, Greece.
6. Treichler, J. R., & Agee, B. G. (1983). A new approach to multipath correction of constant modulus signals. *IEEE Transactions on ASSP*, *31*(2), 459–472.
7. Moir, T. J. (1995). Analysis of an amplitude-locked loop. *Electronics Letters*, *31*(9), 694–695.
8. Pettigrew, A. M., & Moir, T. J. (1993). Coherent demodulation of DSSC without pilot tone using the amplitude-locked loop. *Journal of the Audio Engineering Society*, *41*(12), 998–1007.

Journal: **11277-WIRE** Article No.: **1392** ☐ TYPESET ☐ DISK ☐ LE ☐ CP Disp.:**2013/8/29** Pages: **21** Layout: **Small**

9. Pettigrew, A. M., & Moir, T. J. (1991). Reduction of FM threshold effect by In-band noise cancelling. *Electronics Letters*, *27*(12), 1082–1084.
10. Clark, D. F., & Moir, T. J. (1999). Application of a PLL and ALL noise reduction process in optical sensing systems. *IEEE Transactions on Industrial Electronics*, *44*(1), 136–138.
11. Reed, J. H. (2002). Software radio: A modern approach to radio engineering. Englewood Cliffs: Prentice Hall Professional.

Author Biographies

Thomas J. Moir was born in Dundee Scotland. He was sponsored by GEC Industrial Controls Ltd, Rugby Warwickshire UK from 1976 to 1979 during his B.Sc in control engineering which he was awarded in 1979. In 1983 he received the degree of Ph.D for work on self-tuning filters and controllers. From 1982 to 1983 he was with the Industrial Control unit University of Strathclyde Scotland. From 1983 to 1999 he was a lecturer then senior lecturer at Paisley College/University of Paisley Scotland. Moving to Auckland New-Zealand in 2000, he was with Massey University for 10 years at the Institute of Information and Mathematical Sciences followed by the School of Engineering and Advanced Technology. He moved to AUT University Auckland in 2010 as an associate Professor in the School of Engineering where he works in the area of signal processing and automatic control engineering. He has authored around 100 publication in these fields and is chairman of the Signals and Systems group. He is the holder of one US patent on amplitude-locked loop circuits.

Archibald M. Pettigrew was born in Glasgow,Scotland, in 1945. He received the B.Sc degree in electrical engineering from Glasgow University in 1966 and the M.Sc degree in control engineering from Strathclyde University in 1968. He worked first with Ferranti Ltd., Edinburgh on servo design for a laser range-finding system. From 1973 to 1976 he worked as a linear applications engineer with Signetics, Linlithgow. His two specialities were application of phase-locked loops and the Dolby noise-reduction system. In 1976 he joined Burroughs Ltd.,Glenrothes, as a design manager responsible for the design of Winchester disk-drive systems. He subsequently was responsible for the design and development of a number of magnetic storage products, namely, floppy-disk drives, magnetic card-readers and cassette tape streamers. From 1986 he worked as a consultant engineer/lecturer at the University of Paisley. In 1990 he formed Ampsys Ltd to further develop work in magnetic recording and the amplitude-locked loop. Mr Pettigrew has filed a large number of patents in disk-drive decoding, data-channel design new methods of audio recording and FM noise reduction methods. He is the holder of two John Logie Baird awards for innovation and several UK government awards for research. Since 2011 he has been retired.