On the Linear Representation of GMSK Modulation

Thomas Tsou

tom@tsou.cc

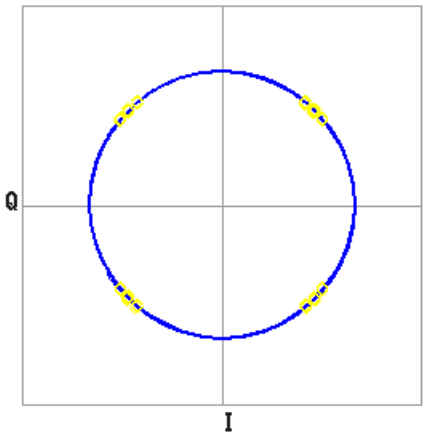


Figure 1: GSM reference signal, phase error 0.29° RMS, 0.85° peak

The GSM standard specifies Gaussian minimum-shift keying (GMSK) as the modulation scheme, which determines the metrics that are used for characterizing signal integrity and performance [1]. In a operational system, such as OpenBTS, there are many factors that affect transmit signal performance prior to reaching the power amplifier: DC offset, IQ imbalance, sampling, and phase noise among others. But, certain factors are hardware related and others are inherent to the design and implementation of the digital transmitter itself; the focus of this chapter is on the latter.

More specifically, this chapter traces signal integrity in the form of 3GPP specified phase error from the hardware output to the core design of the GMSK transmitter. For this process, definitive measurements are made with a GSM specific signal analyzer. We will show that the GSM requirements for GMSK signal quality are not particularly strict and that there is significant margin for improvement beyond minimum compliance values.

0.1 GSM Modulation

A GSM test signal generated by a commercial cellular signal generator (Agilent E4438C) is shown in Figure 1. The receiving and measuring device is a corresponding cellular signal analyzer (E4406A) of the same manufacturer. We note two defining characteristics of the GSM modulation. First, there is the very clear constant envelope and circular phase path. Second, the sampling points (in yellow) form small cluster of three distinct locations; the clustering is the result of intersymbol interference, ISI, of the partial response Gaussian pulse, which will be explained in following sections. Unsurprisingly for calibrated and certified commercial test equipment, measured phase error – the RMS and peak phase deviation compared to an ideal signal – is very low at below 1° and demonstrates an extremely accurate signal.

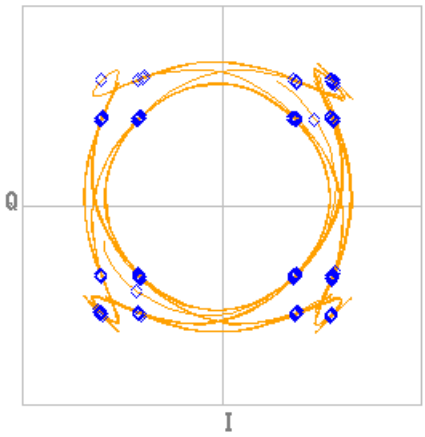


Figure 2: OpenBTS default configuration, phase error 4.98° RMS, 14.95° peak

Figure 2 shows the default output from OpenBTS with USRP1 hardware. In contrast to our reference signal, the OpenBTS output shows non-constant amplitude and irregular phase trajectory. The measured phase error is much higher, but we also note that this signal is still compliant to the GSM specification. The specification 3GPP TS 45.005 [2] states,

> The RMS phase error (difference between the phase error trajectory and its linear regression on the active part of the time slot) shall not be greater than 5 with a maximum peak deviation during the useful part of the burst less than 20.

The RMS measurement is very close to the limit, but the specification tolerates an additional 5 degrees of phase deviation in the peak value. From these observations, we can conclude, first, that the GSM specification is extremely tolerant with regards to phase error and, secondly, that there is substantial room for signal improvement in the OpenBTS transmitter. For the goals presented in this chapter, we seek to reduce the large, distressing gap between reference and observed OpenBTS signals.

0.2 Sampling Effects

The default configuration of OpenBTS runs the transceiver with sampling at 1 complex sample-per-symbol. This factor provides the most immediate source of distortion given that the sampling rate is below Nyquist if we require, for example, a cutoff point of 40 dB attenuation. The Nyquist–Shannon sampling theorem shows that a bandlimited signal can be perfectly reconstructed if sampled at a rate that exceeds twice the bandwidth of the signal. Consequently, default complex sampling at 1 sample-per-symbol – or 2 samples-per symbol when real sampling is considered – accommodates up to a 270.833 kHz signal, which we deem insufficient for accurate reconstruction.

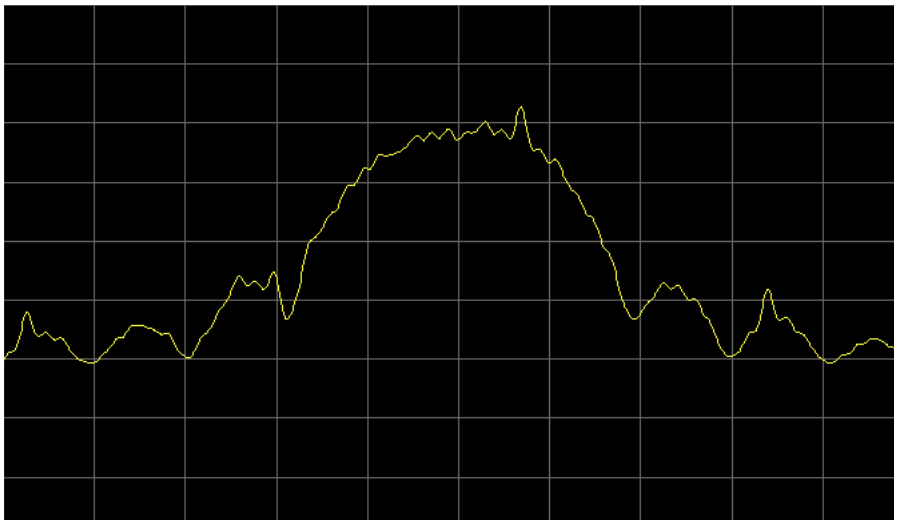


Figure 3: Oversampled OpenBTS spectrum with 8 samples-per-symbol, 1 MHz span

Figures 3 and 4 show the final spectrum of oversampled and default OpenBTS signals respectively. Each capture shows 1 MHz of spectrum with 10 dB increments on the Y-axis. Note that in the oversampled spectrum the occupied bandwidth (roughly 300 kHz at 99% of total signal power) exceeds the available complex sampling bandwidth of the 1 sample-per-symbol case. This bandlimiting is further compounded by downstream filtering by the USRP hardware, which creates additional aliasing effects at the spectrum edges. Consequently, given the observable spectrum distortion, we can conclude that the signal shown in Figure 2 suffers from the effects of insufficient sampling.

Therefore, increasing the sample rate is the initial step to improving transmit signal quality; the same OpenBTS signal sampled at 4 samples-per-symbol is shown in Figure 5. Though 2 samples per symbol may be sufficient from a standalone DSP perspective, these examples are further oversampled to 4 samples-per-symbol for the purpose of reducing potential filter distortion generated upstream on the USRP devices. The CIC (cascaded integrator-comb) filters located on the device apply a sinc shaped frequency response to the input signal, which we can reduce by oversampling.

Given sufficient sampling rate, we now observe a much cleaner phase trajectory and more accurate sampling points. Phase error is now well below the requirement for GSM compliance. Note, though, that there are remaining phase artifacts that were clearly not present in the reference signal of Figure 1 Resolving these phase path effects, which are not a result of sampling, requires a much deeper examination into the design of the OpenBTS modulator. The phase artifacts are a result of the fact the transmitter does not use an exact, direct form GMSK modulator, but instead uses a linearized approximation of GMSK. In order to further pursue the source of these effects, we need to understand the theory and representations of GMSK and the broader class of continuous phase modulations.

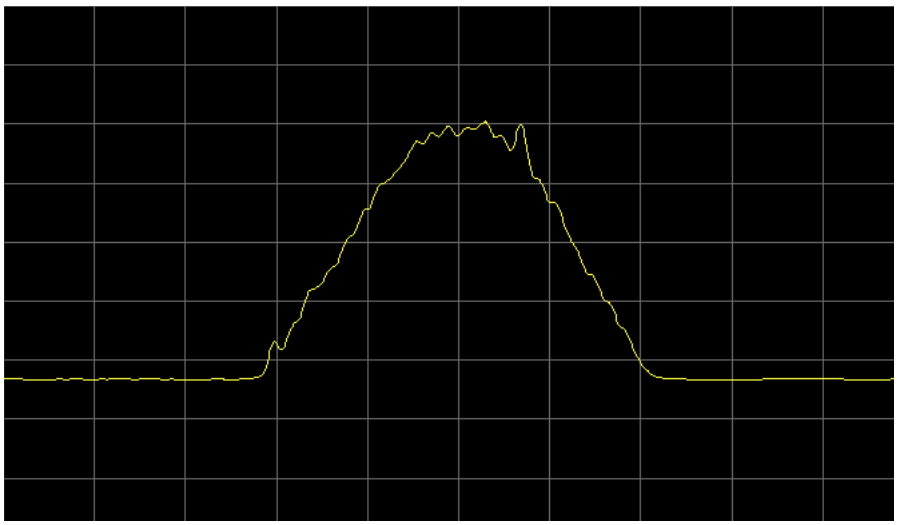


Figure 4: Undersampled OpenBTS spectrum with 1 samples-per-symbol, 1 MHz span

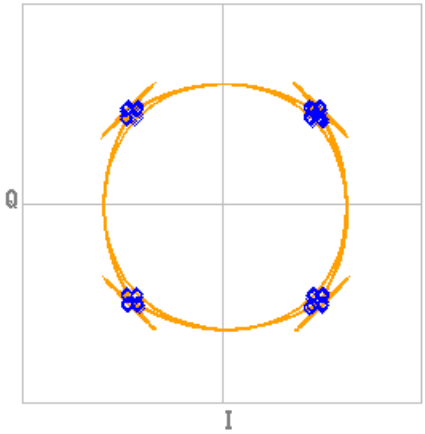


Figure 5: OpenBTS signal, 4 samples per symbol, phase error 1.97° RMS, 5.08° peak

0.3 Continuous Phase Modulation

Continuous Phase Modulation (CPM) belongs to a class of modulation techniques where the phase is constrained to be continuous [3]. This characteristic results in a non-linear system with memory. In the general case, amplitude need not be constrained, however, we assume that information symbols affect the phase only and do not consider the more general case of multi-amplitude CPM.

We begin with the following representation of the carrier modulate CPM signal

$$s(t) = exp\ j\left[2\pi f_c t + \phi(t;I) + \phi_0\right]$$

where f_c is the modulated carrier frequency and $\phi(t;I)$ represents the time varying carrier phase. ϕ_0 is the arbitrary initial phase of the carrier, which we will ignore from this point on. The time varying phase is defined by

$$\phi(t;I) = 2\pi \sum_{k=-\infty}^{n} I_k h_k q(t-kT), \qquad nT \le t \le (n+1)T$$

where I_k is a sequence of information symbols and h_k is a sequence of modulating indices. $q(t)$ is the normalized waveform shape, which is generally represented as the integral of some pulse shape.

$$q(t) = \int_0^t g(\tau)d\tau$$

For simplicity and applicability to GSM signals, we restrict the information sequence I_k to binary symbols $|I_k| = 1$. Furthermore, the modulation index can be fixed for all symbols at a value of $h = \frac{1}{2}$. The use of the latter yields a special form of CPM called Minimum Shift Keying (MSK). We can rewrite the phase of MSK as

$$\begin{aligned}\phi(t;I) &= \frac{1}{2}\pi \sum_{k=-\infty}^{n-1} I_k + \pi I_n q(t-kT), \qquad nT \le t \le (n+1)T \\ &= \theta_n + \frac{1}{2}\pi I_n \left(\frac{t-nT}{T}\right)\end{aligned}$$

Here θ_n is the accumulation of previous phase shifts up to symbol n. Consequently, we have a representation where the current phase at time t is represented by summed previous phase shifts and a positive or negative $\frac{\pi}{2}$ shift for the current symbol. From this point forward, we primarily consider the MSK case where $h = \frac{1}{2}$.

0.4 Laurent Decomposition of CPM Signals

As an alternate representation, CPM can be constructed as a linear combination of a finite sequence of pulses. This second approach provides an additional method that can be used in

the modulation or demodulation process. We will discuss in later implementation sections as to why such an approach is desirable. But, for now, the linearized approach has the benefit of simpler and more convenient implementation and implications towards modulations used in more recent standards. Originally described by Pierre Laurent in 1986 [4], we now provide the linear representation of CPM and follow with a representation for 0.30 BT GMSK that is used for GSM.

We begin with the previous low-pass representation and ignoring the initial phase offset.

$$s(t) = exp\ j\left[\phi(t;I)\right]$$

with the time varying phase described by

$$\phi(t;I) = 2\pi \sum_{k=-\infty}^{n} I_k h_k q(t-kT), \qquad nT \le t \le (n+1)T$$

which we can use to rewrite the signal as

$$e^{\phi(t;I)} = exp\left(j\pi h \sum_{k=-\infty}^{n-L}\right) \prod_{k=0}^{L-1} \exp\left[j2\pi h I_{n-k} q\left(t-(n-k)\,T\right)\right]$$

Laurent then introduces the *generalized phase pulse function* that is derived from the phase shift of the signal. This important function has non-zero values from $0 \le t \le 2LT$ and is specified by

$$s_0(t) = \frac{sin\left(\psi(t)\right)}{sin\pi h}$$

where

$$\begin{aligned} \psi(t) &= 2\pi h q(t) && t < LT \\ \psi(t) &= \pi h - 2\pi h q(t-LT) && LT \le t \end{aligned}$$

This function serves as the basis for constructing the pulse series that will constitute the complete linear representation of the signal. The pulses $c_k(t)$ for $0 \le k \le 2^{L-1}$ are defined by

$$c_k(t) = s_0(t) \prod_{n=1}^{L-1} s_0\left(t + (n + La_{k,n})T\right), \qquad 0 \le t \le Tmin_n\left[L(2-a_{k,n}) - n\right]$$

Each pulse is weighted by a complex coefficient where

$$A_{k,n} = \sum_{m=-\infty}^{n} I_m - \sum_{m=1}^{L-1} I_{n-m} a_{k,m}$$

$$k = \sum_{m=1}^{L-1} 2^{m-1} a_{k,m}, \qquad k = 0, 1, 2, \cdots, 2^{L-1} - 1$$

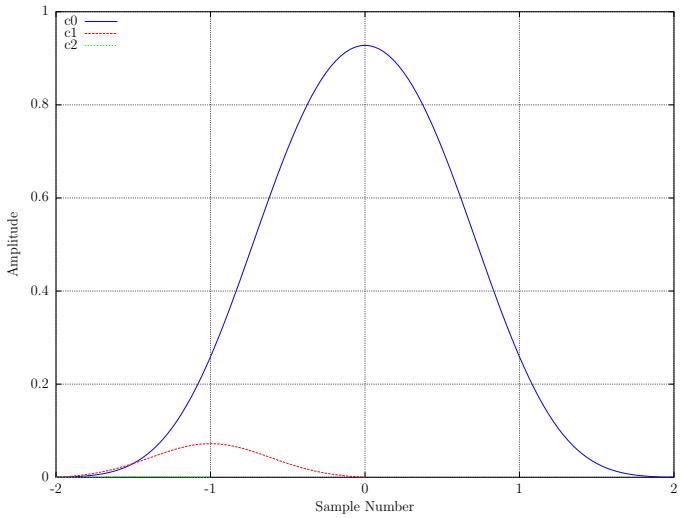


Figure 6: Laurent C0, C1, and C2 pulses for $BT = 0.30$

$$a_{k,m} = 0 \text{ or } 1$$

Given the above phase pulses, substituting and reorganization of terms in the signal function yields the final result

$$e^{\phi(t;I)} = \sum_{n} \sum_{k=0}^{2^{L-1}-1} e^{j\pi h A_{k,n}} c_k(t - nT)$$

which is the CPM signal represented by finite sum of weighted pulses.

0.4.1 GMSK Laurent Decomposition for L=3

As an supplementary example to the mathematical representation, the effects of Laurent decomposition can be shown through graphical comparison. We examine the case for L = 3, which can be easily shown to be sufficient for GMSK with a 0.30 bandwidth time product. The first three generated pulses (named C0, C1, and C2) for $BT = 0.30$ are shown in Figure 6

From the preceding analysis, we can construct our CPM signal with a finite number of pulses, but Figure 6 shows that the practical number of pulses may be an even smaller number than the mathematical representation specifies. In this case the plot shows that only the first two pulses are of any significant magnitude, while the third pulse, C2, is barely visible. Furthermore, the primary pulse, C0, is overwhelmingly dominant. In fact, the C0 pulse alone is sufficient for a linearized approximation of the GMSK modulation, which is used by OpenBTS and measured in Figure 5. But, as we observed, there are remaining artifacts due to the missing C1 pulse, which, though small, are not insignificant.

For comparison purposes, we can also examine the generated Laurent pulses for varying

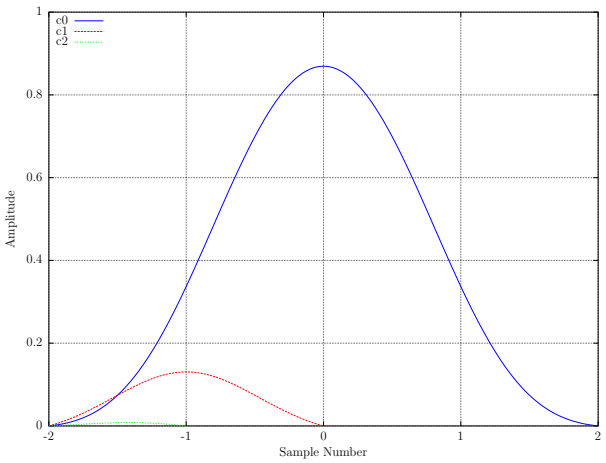


Figure 7: Laurent C0, C1, and C2 pulses for $BT = 0.20$

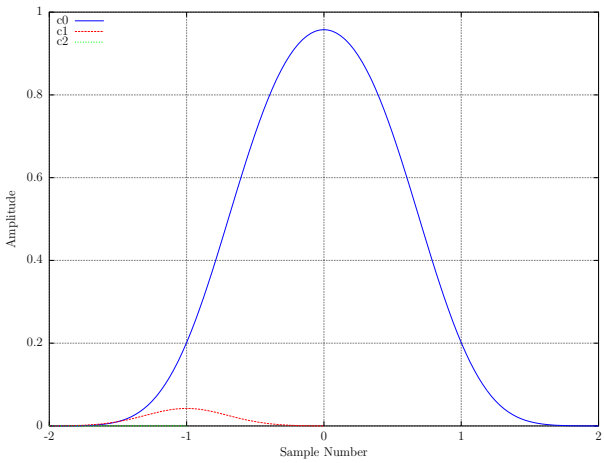


Figure 8: Laurent C0, C1, and C2 pulses for $BT = 0.40$

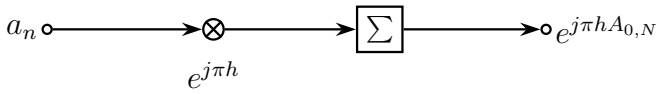


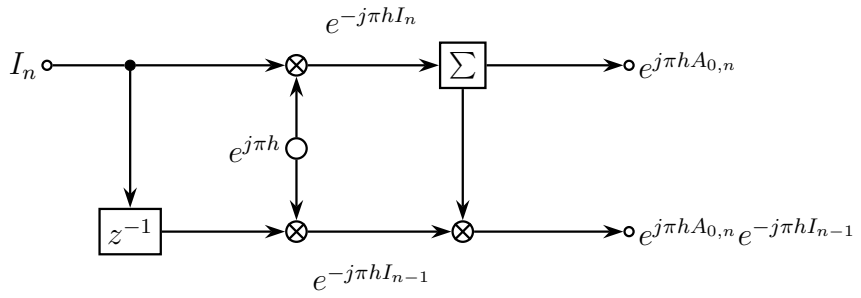
Figure 9: Linearized GMSK symbol mapper with single pulse

widths of the original partial response Gaussian shape. Figures 7 and 8 show the computed pulses for $BT = 0.20$ and $BT = 0.40$ respectively. In all cases we find negligible significance of pulses beyond C0 and C1. We also find an increase in significance of the C1 pulse as the BT product, and ISI, decreases. This is expected since as the BT product approaches infinity, the partial response GMSK signal becomes a normal full response MSK signal, which can be fully represented by linear representation equivalent to offset-QPSK with a half-sinusoidal pulse shape.

0.5 Implementation

Given the linearized representation of CPM, we can now proceed to the actual software implementation. In this section, we directly translate the analytical work of Laurent into a fully operational real time GMSK modulator. The initial OpenBTS design, which uses single C0 pulse implementation is shown in Figure 9. The implementation consists of an input bit sequence mapped to either a positive or negative phase shift of $\pi h = \frac{\pi}{2}$. The phase shift is then accumulated and sent to the pulse shaping filter. Note that GSM employs differential encoding, which is not shown.

In order to add the second C1 pulse, Laurent's decomposition translates to the structure of



$$s(t) = \sum_{n=-\infty}^{\infty} e^{j\pi hA_{0,n}}[c_0(t-nT) + e^{-j\pi hI_{n-1}}c_1(t-nT)]$$

$$A_{0,n} = \sum_{m=-\infty}^{n} I_m$$

$$A_{1,n} = A_{0,n} - I_{n-1}$$

$$I_n = 1, -1$$

Figure 10: Linearized GMSK symbol mapper with C0 and C1 pulses

Figure 10. While there is an expanded structure to accommodate the effect of one delayed signal, the operations for MSK, $h = \frac{1}{2}$, are simply bit inversions in the in-phase or quadrature channels. Filtering and summing with the C0 and C1 pulse shapes can be implemented with a structure that resembles two conjoined tapped delay lines as shown in Figure 11. The tap lengths in this case represent computed pulse filters at 4 samples-per-symbol.

0.6 Results and Discussion

With the advantages of the linearized approach, our pulse simulations showed that we can easily ignore all pulses except for C0 and C1. That led us to the construction given in Figures 11 and 10. We also know that we can only use the C0 pulse for a very simple implementation, but then suffer from the phase trajectory artifacts shown at the beginning of this chapter. The USRP1 output of our final construction is shown in Figure 14 alongside the original single pulse implementation at 1 and 4 samples per symbol in Figure 12 and Figure 13 respectively.

Given the modified GMSK modulator of Figure 14, we can clearly observe the absence of phase artifacts present in the original modulator, Figure 13. Obviously, both cases show substantial improvements from the sub-Nyquist samples case of Figure 2. While modifying

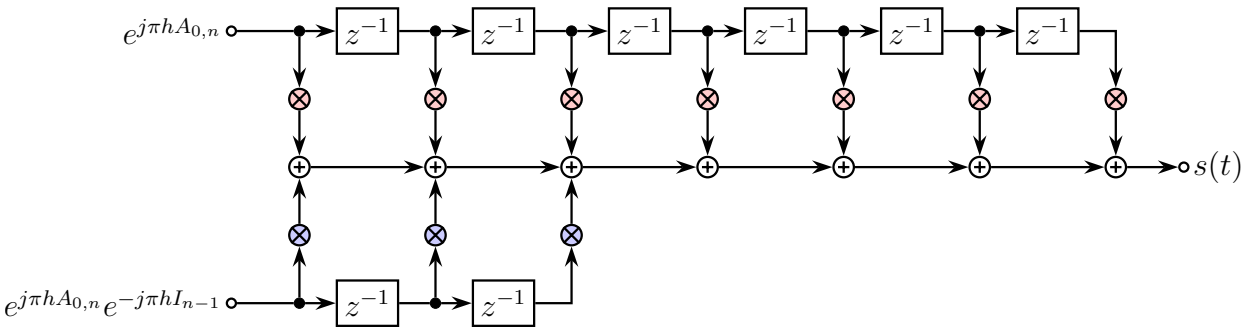


Figure 11: Combined tapped delay line for linearized GMSK pulses C0 and C1

the OpenBTS modulator shows clearly visible improvements in the magnitude error near the sampling points, the improvement in measured phase error is quite modest. To find out why, we can measure the same modulator implementation through an USRP2, which is displayed in Figure 15.

The USRP2 shows superior phase accuracy – recall that the GSM compliance phase error is 5° RMS and 20° peak. The summary of various modulator and hardware combinations discussed in this chapter are shown in Table 1. In the USRP2 case, the primary difference between USRP1 and USRP2 measurements are due to the availability of DC offset and IQ imbalance corrections on the USRP2. Both combinations use the same WBX daughterboard from Ettus Research, but only the USRP2 supports automatic daughterboard calibration. Consequently, we can conclude that, to a large extent, we are no longer primarily observing GMSK modulator irregularities, but the effects of uncompensated RF hardware differences.

The linear modulation approach described in this chapter, at first glance, seems unnecessary and cumbersome given that a basic GMSK implementation with a pulse shaped bit sequence driving a phase accumulator is not particularly complex. There are, however, three significant advantages of the linearized approach. First, the modulator can be implemented using only standard multiply-accumulate operations. This contrasts to the direct implementation of MSK that drives a frequency modulator using trigonometric methods, which may be difficult to implement or optimize depending on the processor architecture. Second, while GMSK modulation is constant amplitude, the TDMA based GSM signal is not. Time slots in GSM are separated by guard intervals with specified amplitude ramp-up and ramp-down requirements [5]. Consequently, the implementation does need to consider amplitude changes, and the linear approach can provides a simpler implementation. Finally, there are strong implications for the linearized approach because of developments in later specifications. For example, the linearized Gaussian pulse shape is explicitly specified for use in 8-PSK EDGE [1]. Furthermore, the same pulse shape is used again with a modified QPSK modulation in the more recent Release 9 specification for orthogonal subchannels and

Table 1: Measured Phase Error

	Samples-per-symbol	RMS	Peak
3GPP TS 45.005	N/A	5.00	20.00
USRP1, C0	1	4.89	14.25
USRP1, C0	4	1.92	5.44
USRP1, C0 and C1	4	1.66	4.53
USRP2, C0 and C1	4	1.14	2.97

voice, VAMOS [6]. Given these concerns, the linearized approach to GMSK turns out to be very reasonable for GSM targeted implementations.

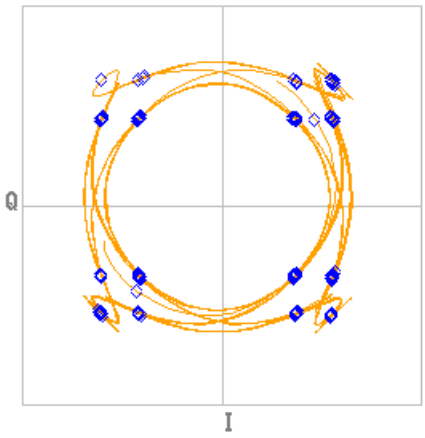


Figure 12: OpenBTS–USRP1 phase error, 1 sample per symbol, C0 pulse, 4.98° RMS, 14.95° peak

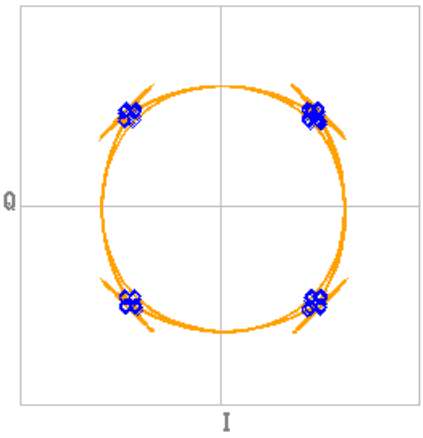


Figure 13: OpenBTS–USRP1 phase error, 4 samples per symbol, C0 pulse, 1.92° RMS, 5.44° peak

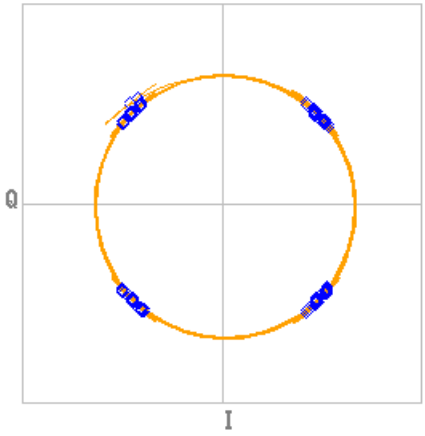


Figure 14: OpenBTS–USRP1 phase error, 4 samples per symbol, C0 and C1 pulses, 1.66° RMS, 4.53° peak

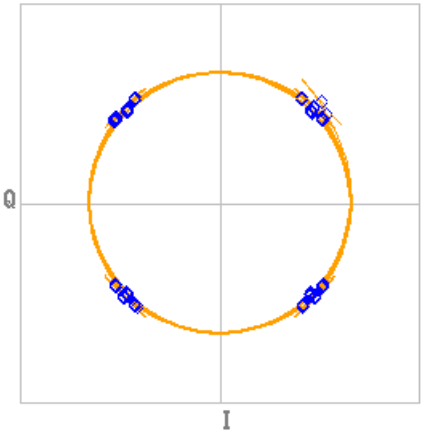


Figure 15: OpenBTS–USRP2 phase error, 4 samples per symbol, C0 and C1 pulses, 1.14° RMS, 2.97° peak

Bibliography

[1] 3GPP, "Modulation," 3rd Generation Partnership Project (3GPP), TS 45.004, Sep. 2008.

[2] ——, "Radio transmission and reception," 3rd Generation Partnership Project (3GPP), TS 45.005, Sep. 2008.

[3] J. Proakis, *Digital Communications*. McGraw-Hill, 1995.

[4] P. Laurent, "Exact and Approximate Construction of Digital Phase Modulations by Superposition of Amplitude Modulated Pulses (AMP)," *Communications, IEEE Transactions on*, vol. 34, no. 2, pp. 150 – 160, Feb 1986.

[5] 3GPP, "Physical layer on the radio path; General description," 3rd Generation Partnership Project (3GPP), TS 45.001, Sep. 2008.

[6] M. Saily, G. Sebire, E. Riddington, *GSM/Edge: Evolution and Performance*. Wiley, 2011.