INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI[®]

Modulation and Frequency Synthesis for Wireless Digital Radio

by

Walter T. Bax B. Eng., M. Eng.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Ottawa-Carleton Institute of Electrical Engineering
Department of Electronics
Carleton University
Ottawa, Canada

© Copyright October 1999

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-48341-X

The undersigned hereby recommend to the Faculty of Graduate Studies and Research
acceptance of the thesis

Modulation and Frequency Synthesis for Wireless Digital Radio

submitted by

Walter T. Bax B. Eng., M. Eng.

in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Chair, Department of Electronics

Thesis Supervisor

External Examiner

Carleton University
October 1999

Abstract

A new wideband modulator architecture, that is suitable for continuous-phase constant-envelope modulation schemes, is presented in this thesis. The technique is based on direct modulation of a high resolution $\Delta\Sigma$ frequency discriminator based synthesizer to produce the modulated RF signal without up-conversion. The advantage of this architecture is that it does not require mixers or D/A converters to generate the In-phase and Quadrature signals as in conventional GMSK modulators. This eliminates many of the analog problems associated with mixing and filtering and results in an architecture suitable for monolithic integration.

A high modulation data rate is possible, without sacrificing phase noise performance, through digital equalization of the synthesizer closed-loop response. A digital GMSK transmit filter pre-shapes the data symbols and additionally compensates for the synthesizer closed-loop response. This permits the use of a narrower synthesizer bandwidth to attenuate the $\Delta\Sigma$ quantization noise, while equalization effectively widens the modulation bandwidth to handle high data rates. The digital equalization filter adds little complexity to the transmitter architecture, since it is combined with the Gaussian data filter. Matching between the transmit filter and the synthesizer closed-loop response is not an issue since the loop parameters are digitally defined and are therefore predictable.

An experimental GSM modulator operating at 2GHz was developed to validate the suitability of the architecture for use in wireless communication systems. It makes use of a custom BiCMOS $\Delta\Sigma$ frequency discriminator chip that is a key component in the modulator architecture.

Acknowledgments

I would like to express my gratitude to Robert Hadaway and Dr. Peter Schvan of Nortel's Technology Access and Applications group for their technical and financial support over the course of this research. Thanks also to microsurgeon Robin Collins and countless others in manufacturing.

The generous financial support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Telecommunications Research Institute of Ontario (TRIO) is gratefully acknowledged.

The endless support of Nagui Mikhail over the years and the assistance of Jorge Aguirre during hardware testing made the final results possible.

The final word goes to Professor Miles Copeland. Thanks for overseeing my efforts and guiding me in the right direction.

Table of Contents

List of Figures	viii
List of Tables	xiii
List of Symbols and Abbreviations	xiv
1 Introduction	1
1.1 Contributions	.1
1.2 Thesis Outline	.3
2 Wireless Digital Radio	4
2.1 Digital Modulation Techniques	.5
2.1.1 Linear Modulation	.5
2.1.1.1 Binary Phase Shift Keying	.6
2.1.1.2 Quadrature Phase Shift Keying	.9
2.1.2 Constant envelope modulation	.11
2.1.2.1 Binary Frequency Shift Keying	.12
2.1.2.2 Minimum Shift Keying	.14
2.1.2.3 Gaussian Minimum Shift Keying	.17
2.2 Constant Envelope Modulators	.20
2.2.1 Mixer Based	.21
2.2.2 Direct Modulation	.22
2.2.3 Indirect Modulation	.24
2.2.3.1 Narrow Band Indirect Modulation	.24
2.2.3.2 Offset Phase-Locked Loop	.26
2.2.3.3 Wideband Indirect Modulation	.29
3 Continuous-Phase Modulation Using a $\Delta\Sigma$FD Based Synthesizer	33
3.1 Wideband Modulator Architecture	.33
3.2 Modelling	.37
3.2.1 $\Delta\Sigma$ Frequency Discriminator Model	.38

3.2.2	Synthesizer Model	48
3.3	Design Parameters	53
3.3.1	PLL Order	53
3.3.2	Reference Frequency	54
3.3.3	Stability	55
3.4	Equalization of Synthesizer Closed-Loop Response	61
3.4.1	Transmit Filter.	61
3.4.2	Dynamic Range Constraints	63
3.4.2.1	$\Delta\Sigma$ Frequency Discriminator Overload	65
3.4.2.2	Digital $\Delta\Sigma$ Modulator Range	67
3.4.3	Modulation Bandwidth Limitations.	70
3.4.3.1	Effect of Sampling Frequency	71
3.4.3.2	Synthesizer Loop Bandwidth	74
3.4.4	Effect of Mismatch	74
3.4.4.1	Loop Stability	75
3.4.4.2	Open-Loop Gain Error	75
3.5	Gaussian Pulse Shaping.	77
3.6	Digital $\Delta\Sigma$ Modulator	78
4	A 2.5GHz BiCMOS $\Delta\Sigma$ Frequency Discriminator	82
4.1	$\Delta\Sigma$ Frequency Discriminator Architecture	83
4.1.1	Non-linear Effects.	86
4.1.2	Enhancing the Input Sensitivity.	91
4.1.3	Loop Stability	96
4.1.4	Acquisition	103
4.1.5	Achievable Signal-to-Noise Ratio	107
4.1.6	Influence of Circuit Parameters	112
4.2	BiCMOS $\Delta\Sigma$ Frequency Discriminator Chip	115
4.2.1	High Speed, Low Power Design Techniques	116
4.2.2	Multi-Modulus Divider with Low Delay	120
4.2.3	Phase-Frequency Detector	129
4.2.4	Charge Pump	134
4.2.5	Quantizer	138
4.2.6	Noise Calculations	141
4.2.7	Mixed Signal Design and Layout Techniques	149
4.2.8	Measured Results	153
5	Modulator Design and Implementation	160
5.1	GSM Design Example	160
5.2	Mixed-Signal Synthesizer Blocks.	165
5.2.1	Digital Signal Processor.	165
5.2.2	Digital-to-Analog Converter	175
5.2.3	Analog Integrator	176
5.2.4	Voltage-Controlled Oscillator.	180

5.3	Modulation Data Path	181
5.3.1	Digital Transmit Filter	182
5.3.2	Digital MASH $\Delta\Sigma$ Modulator	184
5.4	Noise Analysis	186
5.5	Modulator Performance	193
5.5.1	GMSK Transmitter Performance	193
5.5.2	Synthesizer Performance	199
6	Conclusion	206
6.1	Future Research	208
	References	209
	Appendix A GMSK Modulator	214
	Appendix B $\Delta\Sigma$ Frequency Discriminator	217

List of Figures

2.1	Unfiltered PSK modulated carrier.	6
2.2	Raised cosine impulse response.	7
2.3	Power spectral density of an unfiltered and raised cosine filtered ($\alpha=0.5$) BPSK signal.	8
2.4	BPSK constellation diagram: (a) unfiltered and (b) filtered.	9
2.5	QPSK constellations with different signal sets.	10
2.6	Power spectral density of an unfiltered and raised cosine filtered ($\alpha=0.5$) QPSK signal.	11
2.7	Phase trajectory of an unfiltered MSK signal.	15
2.8	Constellation of an MSK signal.	15
2.9	Power spectral density of MSK and QPSK modulated signals.	16
2.10	Impulse response of Gaussian filtered and unfiltered MSK signals.	18
2.11	Power spectral density of GMSK signals with various bandwidths.	18
2.12	Phase trellis of MSK and GMSK signals.	19
2.13	Constellation of MSK and GMSK signals.	20
2.14	Block diagram of a quadrature amplitude modulator (QAM).	21
2.15	DECT open-loop modulator.	23
2.16	$\Delta\Sigma$ fractional-N synthesizer.	25
2.17	Modulator with time-varying reference frequency.	27
2.18	Block diagram of an offset phase-locked loop (OPLL).	28
2.19	Block diagram of a wideband modulator using a digital equalizer and pulse shaping filter.	29
3.1	$\Delta\Sigma$ frequency discriminator based synthesizer [Bax95].	34
3.2	Equalized direct modulation of a $\Delta\Sigma$ FD based synthesizer.	35
3.3	First-order $\Delta\Sigma$ frequency discriminator (a) model and (b) realization.	39
3.4	Digital phase-frequency detector timing diagram.	40
3.5	Timing diagram of first-order $\Delta\Sigma$ frequency discriminator.	41
3.6	Digital multi-modulus divider model.	43
3.7	First-order $\Delta\Sigma$ frequency discriminator model.	44
3.8	Equivalent (a) multi-loop and (b) single-loop second-order $\Delta\Sigma$ modulator structures.	45

3.9	Single-loop second-order $\Delta\Sigma$ frequency discriminator.	46
3.10	Second-order $\Delta\Sigma$ frequency discriminator model.	47
3.11	Linear equivalent noise model of a second-order $\Delta\Sigma$ frequency discriminator.	48
3.12	Linearized equivalent model of $\Delta\Sigma$ FD based synthesizer.	49
3.13	Modulation data path.. . . .	50
3.14	Second-order synthesizer open-loop Bode plot.	57
3.15	Comparison between discrete-time (13MHz sampling frequency) and continuous-time synthesizer open-loop transfer functions.. . . .	60
3.16	Transmit filter composed of Gaussian and equalizer responses.	62
3.17	Baseband filter response with (a) unrestricted and (b) restricted modulation bandwidths.	62
3.18	Block diagram of the $\Delta\Sigma$ FD based GMSK modulator.	64
3.19	Reducing the D/A dynamic range through remodulation.	65
3.20	Single-stage $\Delta\Sigma$ modulator.	68
3.21	Noise power of a first-order $\Delta\Sigma$ modulator (OSR=8).	69
3.22	Multi-stage (MASH) $\Delta\Sigma$ modulator.	69
3.23	Block diagram of second-order $\Delta\Sigma$ frequency discriminator.	71
3.24	Effect of $\Delta\Sigma$ FD divider modulus n on PFD peak phase error.	73
3.25	Modulation data path.. . . .	76
3.26	Misshaped modulation transfer function due to open-loop gain error.	76
3.27	GMSK baseband modulation bandwidth with varying (a) filter bandwidth BT and (b) symbol rate.. . . .	78
3.28	Digital second-order MASH $\Delta\Sigma$ modulator.	80
4.1	Single-loop, second-order $\Delta\Sigma$ frequency discriminator.	83
4.2	Second-order single-loop $\Delta\Sigma$ frequency discriminator model.	83
4.3	Second-order frequency discriminator: (a) noise transfer function and (b) pole-zero plot.. . . .	84
4.4	Comparison between Nyquist, oversampled and $\Delta\Sigma$ noise shaped quantization.	85
4.5	Non-linear SIMULINK, model used for time-domain simulation.	86
4.6	Noise power of a first-order $\Delta\Sigma$ modulator (OSR=8).	87
4.7	Deadzone effect in a second-order $\Delta\Sigma$ modulator with integrator leakage (gain=64) and OSR=64.	89
4.8	Modified single-loop $\Delta\Sigma$ frequency discriminator block diagram.	93
4.9	Digital matched filter response for GSM modulation.	95
4.10	Filtered GSM eye diagram from (a) original and (b) modified $\Delta\Sigma$ FD output.	95
4.11	Simplified second-order $\Delta\Sigma$ frequency discriminator linear stability model.	96
4.12	Linear signal-dependent model of a 1-bit quantizer.	97
4.13	Signal dependent gain of a 1-bit quantizer (modulus $n=142$).	98
4.14	Root-locus plot of the linear $\Delta\Sigma$ FD model.. . . .	99
4.15	$\Delta\Sigma$ FD signal transfer function for various quantizer gains.	100
4.16	$\Delta\Sigma$ FD noise transfer function for various quantizer gains.. . . .	101
4.17	$\Delta\Sigma$ FD state space diagram for (a) low frequency, (b) midband frequency and (c) high frequency RF input signals.	103
4.18	$\Delta\Sigma$ FD initial acquisition after power-up.	105

4.19	$\Delta\Sigma$ FD acquisition following an input frequency step.	106
4.20	Signal-to-noise ratio of second-order frequency discriminator (BW=200KHz).	108
4.21	Effect of oversampling ratio on SNR of second-order frequency discriminator with BW=200KHz.. . . .	109
4.22	Effect of quantizer resolution on SNR of second-order frequency discriminator with BW=200KHz.. . . .	110
4.23	Effect of divider fractional- δ on SNR of second-order frequency discriminator with BW=200KHz.. . . .	111
4.24	Effect of PFD deadzone on SNR of second-order frequency discriminator with BW=200KHz.	113
4.25	Simplified single-ended charge pump.	113
4.26	Transit frequency of a 1X bipolar device ($A_E=0.8 \times 4.0 \mu\text{m}$) [Hada91].	117
4.27	A typical ECL/CML logic gate.	118
4.28	Complex ECL/CML logic gate.	120
4.29	Block diagram of low delay multi-modulus divider.	121
4.30	State diagram of low delay multi-modulus divider.	122
4.31	RF buffer with level shifter.. . . .	123
4.32	RF input buffer gain and bandwidth.	123
4.33	Differential 4/5 dual-modulus divider.	125
4.34	Dual-modulus divider timing diagram.	125
4.35	Multi-modulus divider timing diagram.	126
4.36	Multi-modulus divider setup time.	128
4.37	Phase-frequency detector with asynchronous reset.	130
4.38	Differential PFD flip-flop with asynchronous reset.	131
4.39	Phase-frequency detector timing diagram.	132
4.40	Differential phase-frequency detector transfer function.	133
4.41	Simplified representation of a differential charge pump.. . . .	134
4.42	Simplified schematic of differential charge pump with active feedback.. . . .	136
4.43	Differential charge pump timing diagram.	137
4.44	Differential charge pump linearity and output range.	138
4.45	Differential 1-bit quantizer with BiCMOS input buffer/comparator.. . . .	139
4.46	Differential 1-bit quantizer DC transfer characteristic.. . . .	140
4.47	Second-order single-loop $\Delta\Sigma$ frequency discriminator model.	141
4.48	First-order mapping of voltage noise to timing jitter.	143
4.49	Differential PFD output timing jitter.	145
4.50	$\Delta\Sigma$ FD output referred frequency noise spectral density.	146
4.51	Layout plot of high-speed differential 4/5 dual-modulus divider.	151
4.52	Layout plot of differential BiCMOS charge pump.. . . .	153
4.53	Photomicrograph of BiCMOS $\Delta\Sigma$ frequency discriminator.	154
4.54	Output spectrum of $\Delta\Sigma$ FD with DC input (unmodulated carrier).	157
4.55	Output spectrum of measured $\Delta\Sigma$ FD bitstream with 100KHz single-tone FM modulated carrier.. . . .	158
4.56	In-band view of measured $\Delta\Sigma$ FD output spectrum with 100KHz single-tone FM modulated carrier.	159

5.1	Block diagram of the $\Delta\Sigma$ FD based GMSK modulator.	161
5.2	Open-loop transfer function of $\Delta\Sigma$ FD based synthesizer.	164
5.3	FPGA design flow.	166
5.4	Mapping the (a) direct form FIR filter into (b) a ROM based FIR filter.	168
5.5	Equivalent filter structures: (a) smaller ROM size with more adder levels or (b) larger ROM size with less adder levels.	169
5.6	Quantization noise filter: (a) ideal infinite impulse response and (b) scaled and quantized finite impulse response.	171
5.7	Frequency response of ideal Butterworth IIR filter and approximate FIR filter.	172
5.8	Digital synthesizer loop filter employing saturation arithmetic and detection.	173
5.9	Reducing the D/A dynamic range requirement through $\Delta\Sigma$ remodulation.	174
5.10	Continuous-time integrator with variable gain and negative output clamp.	177
5.11	Continuous-time integrator response with non-ideal op-amp ($A_o=87$ dB, UGBW=2MHz).	179
5.12	Open-loop phase noise of Z-COMM model V613ME04 VCO.	181
5.13	Digital modulation data path.	182
5.14	GSM Inter-symbol interference between (a) individual symbols and (b) the combined effect on a symbol sequence.	183
5.15	Digital second-order MASH $\Delta\Sigma$ modulator block diagram.	184
5.16	Discontinuities introduced by splicing finite length $\Delta\Sigma$ modulator outputs.	186
5.17	Noise model for GMSK modulator using linear $\Delta\Sigma$ FD model.	187
5.18	Equivalent block diagram of synthesizer noise sources.	188
5.19	Simulated phase noise of GMSK modulator.	190
5.20	Time-domain modulator output signal composed of equivalent GMSK phase modulation and synthesizer phase noise.	191
5.21	Output spectrum of (a) unmodulated carrier and (b) GMSK modulated carrier.	192
5.22	Output power spectral density for an (a) ideal GSM modulated carrier and (b) simulated GMSK modulator with optimal parameter set.	194
5.23	Measured output power spectrum of GMSK modulator with a 1.8655 GHz carrier frequency.	195
5.24	Vector modulation analyzer test set up.	195
5.25	Reference and measured constellations of GSM modulated carrier ($BT=0.3$).	196
5.26	Simulated (a) and measured (b) I and Q eye diagrams with 0% open-loop gain error.	197
5.27	Simulated (a) and measured (b) I and Q eye diagrams with +20% open-loop gain error.	198
5.28	Simulated (a) and measured (b) I and Q eye diagrams with -20% open-loop gain error.	198
5.29	Measured synthesizer output spectrum.	200
5.30	Measured synthesizer spurious noise.	201
5.31	Simulated phase noise of GMSK modulator using <i>measured</i> $\Delta\Sigma$ FD noise.	202
5.32	Measured synthesizer phase noise.	203
5.33	Simulated synthesizer switching speed for an input frequency step.	204
5.34	Measured synthesizer switching speed for a 6MHz frequency step.	205

A.1	RF and analog section schematic.	215
A.2	Digital signal processor schematic.. . . .	216
B.1	BiCMOS $\Delta\Sigma$ FD chip bonding diagram.	218

List of Tables

3.1	Effect of open-loop gain error on loop parameters.	75
4.1	Divider modulus range for various quantizer resolutions.	90
4.2	$\Delta\Sigma$ FD functional specification for GSM modulation.	116
4.3	Differential charge pump operating modes.	135
4.4	Ideal $\Delta\Sigma$ FD input sensitivity.	148
4.5	Optimal $\Delta\Sigma$ FD input sensitivity.	149
4.6	BiCMOS $\Delta\Sigma$ FD chip DC test results.	155
4.7	BiCMOS $\Delta\Sigma$ FD chip AC test results.. . . .	155
5.1	Synthesizer loop parameters for GSM modulation.	163
5.2	Digital Butterworth filter parameters.	167
5.3	256 tap FIR filter partitioning.	170
5.4	Influence of synthesizer parameters on individual noise sources.	189

List of Symbols and Abbreviations

A/D	analog to digital converter
ASIC	application specific integrated circuit
BER	bit error rate
BFSK	binary frequency shift keying
BPSK	binary phase shift keying
BT	normalized Gaussian filter bandwidth
BW	bandwidth
CMOS	complementary metal oxide semiconductor
CP	charge pump
D/A	digital to analog converter
DECT	digital enhanced cordless telecommunications standard
DMD	dual-modulus divider
$\Delta\Sigma$ FD	$\Delta\Sigma$ frequency discriminator
DSP	digital signal processor
f_r	reference frequency
f_T	bipolar transistor transit frequency
FIR	finite impulse response
FM	frequency modulation
FPGA	field programmable gate array
GFSK	Gaussian frequency shift keying
GMSK	Gaussian minimum shift keying
GSM	global system for mobile communication
IC	integrated circuit
IF	intermediate frequency
IIR	infinite impulse response
ISI	inter-symbol interference
J_K	bipolar transistor critical current density
K	open-loop gain
K_{CP}	charge pump gain
K_ϕ	phase detector gain
K_i	integrator gain

K_q	quantizer gain
K_v	voltage-controlled oscillator sensitivity
$\mathcal{L}(f)$	single-sideband phase noise spectral density
LO	local oscillator
MASH	cascaded $\Delta\Sigma$ modulator
MSE	mean square error
MSK	minimum shift keying
OPLL	offset phase-locked loop
OQPSK	offset quadrature phase shift keying
OSR	oversampling ratio
PA	power amplifier
PCB	printed circuit board
PFD	phase-frequency detector
PLL	phase-locked loop
PSRR	power supply rejection ratio
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
R_b	symbol rate
ROM	read only memory
SNR	signal-to-noise ratio
SSB	single sideband
T_b	symbol duration
T_r	reference period
TDMA	time division multiple access
UGBW	unity gain bandwidth
VCO	voltage-controlled oscillator
VHDL	very large scale hardware description language
VLSI	very large scale integrated circuit
ω_n	natural frequency
ζ	damping factor

Chapter 1

Introduction

The thrust toward low-power radio architectures has led to the use of constant-envelope modulation schemes. This permits the use of non-linear power amplifiers which are much more power efficient than their linear counterparts. The potential power savings of using non-linear power amplifiers can be further exploited by focusing on alternative modulator architectures that are less complex and permit a higher degree of integration.

A new wideband modulator architecture is presented that is suitable for continuous-phase constant-envelope modulation schemes. The technique uses a high resolution synthesizer that can produce the modulated RF signal without up-conversion.

1.1 Contributions

This thesis explores a new GMSK modulator architecture that is suitable for wireless radio applications. It addresses issues in regards to designing low complexity transmitter architectures that are suitable for VLSI integration in complete radio systems. The technique exploits results obtained from earlier research by the author on a new synthesizer architecture that uses a $\Delta\Sigma$ frequency discriminator in the feedback path [Bax94]. While the results in [Bax94] indicated the architecture was feasible for use as a high resolution synthesizer, it was unsuitable for modulation applications. The new GMSK modulator described in this thesis employs a similar synthesizer architecture, heavily modified to match the requirements for transmit applications. The addition of

wideband modulation capability provides a complete solution for the transmit needs of a digital transceiver.

The contributions in this thesis are aimed toward the development of a wideband low-power transmitter architecture for wireless digital radio and are summarized as follows:

- A review of current modulation techniques is presented to illustrate the fundamental advantages and limitations in terms of bandwidth and power efficiency. These results represent the theoretical limits that any particular modulation scheme can achieve in an ideal environment. Arguments that favour the use of constant-envelope MSK modulation schemes are made and some current MSK modulator architectures are described. The feasibility of integrating these modulators structures to produce a wideband low cost transmitter solution is discussed.
- A new wideband constant envelope GMSK modulator that uses a $\Delta\Sigma$ frequency discriminator is presented and addresses the integration problems encountered in previous architectures. The new architecture makes extensive use of digital techniques to ease the analog design constraints, which results in an architecture that is easily integrated in an integrated circuit (IC) technology. The modulator architecture overcomes the limited modulation bandwidth through the use of digital equalization similar to that described in [Perr97]. This technique dramatically extends the maximum modulation bandwidth and allows the parameters that govern the modulation and noise performance to be independently set.
- Models of all the modulator components will be developed, with emphasis on the $\Delta\Sigma$ frequency discriminator, which makes use of a new multi-modulus divider model. Simulations of a complete transmitter model are used to assess the transmitter performance limitations and to reveal the impact of any non-ideal effects.
- Design and implementation issues of the $\Delta\Sigma$ frequency discriminator, which is a fundamental block in the modulator, are presented. The goal is to integrate the entire discriminator in a BiCMOS IC technology, which will permit high frequency operation suitable for digital radio applications in the 2GHz range (e.g. DCS-1800). Since the discriminator lies in the feedback path of the modulator, any error

introduced by it is uncorrected, which implies it must have high performance while operating at high speeds. Single-chip VLSI implementation strategies are introduced that yield the desired bandwidth without incurring a high power consumption.

- An example transmitter design, using the GSM modulation standard, is used to explore the design space of the new architecture. Once the design parameters have been identified, a discrete hardware prototype is used to verify the modulator performance under various conditions.

1.2 Thesis Outline

The thesis is organized in the following manner. Chapter 2 reviews the basic digital modulation schemes and discusses the advantages and disadvantages of each. The use of non-linear, constant-envelope modulation schemes is shown to be suitable for transceiver designs where low power consumption is the primary concern. A new GMSK modulator architecture, using a $\Delta\Sigma$ frequency discriminator based synthesizer first described in [Bax95], is introduced in Chapter 3. It is shown that this architecture has several advantages over more conventional constant-envelope modulators and is more suitable for integration in an IC technology. Design issues and VLSI implementation of the $\Delta\Sigma$ frequency discriminator, a major component in the GMSK modulator, are outlined in Chapter 4. A hardware prototype GMSK modulator that makes use of the single-chip $\Delta\Sigma$ discriminator is covered in Chapter 5. The modulator uses the GSM modulation standard to demonstrate the feasibility of the architecture for wireless digital radio in the 2GHz range. Finally, some concluding remarks on the modulator performance and suggested improvements as well as areas of future research are described in Chapter 6.

Chapter 2

Wireless Digital Radio

Modern mobile communication systems use digital modulation techniques. Advancements in very large scale integration (VLSI) and digital signal processing (DSP) technology have made digital modulation more cost effective than analog transmission systems [Rapp96]. Digital modulation offers many advantages over conventional analog modulation. Some advantages are:

- efficient use of available spectrum through coding and modulation techniques
- easier multiplexing of different forms of data (voice, data and video)
- noise immunity through error correction and channel equalization
- management of complex cellular networks.

The decision to use digital modulation raises the question of what modulation scheme to choose. A suitable digital modulation scheme provides low bit error rates at low received signal-to-noise ratios, performs well in multipath and fading conditions, occupies a minimum bandwidth and is easy and cost effective to implement. No modulation scheme satisfies all these criteria concurrently so one must decide what factors are important for a particular application.

The performance of a particular modulation scheme is often measured in terms of its power and bandwidth efficiency [Rapp96]. Bandwidth efficiency describes the ability of a modulation scheme to accommodate data within a limited bandwidth and is measured in bits per second per Hertz (bps/Hz). Increasing the data rate decreases the pulse width of a digital symbol, which increases the bandwidth of the signal. This consequence applies to

all modulation schemes but some are more efficient than others. Power efficiency describes the transmitted power required for a receiver to preserve the integrity (maximum bit error rate) of the data at low levels. Achieving a certain bit error rate (BER) requires a minimum power which varies according to the modulation scheme used. Power efficiency is often measured as the ratio of energy per bit to noise spectral density (E_b/N_0). In the design of digital communication systems, very often there is a trade-off between power and bandwidth efficiency. An example of this is adding error control, which consumes bandwidth (reduces bandwidth efficiency) but reduces the required received power for a given BER. There are other factors to consider when designing a personal communication system. The cost and complexity of the handset must be minimized to make the system attractive to the subscriber. The performance under various channel impairments (multipath, fading etc.) dictates the type of modulation scheme used.

2.1 Digital Modulation Techniques

Digital modulation techniques may be broadly classified as linear and non-linear. The advantages and disadvantages of both classes are described in the following sections along with some examples of each type.

2.1.1 Linear Modulation

In linear modulation, the amplitude of the transmitted signal varies linearly with the modulating signal. Linear techniques are bandwidth efficient, and hence are attractive for use in systems where there is an increased demand for more users within a limited spectrum. While linear modulation has very good spectral efficiency, the signal must be transmitted using linear RF amplifiers which have very poor power efficiency (i.e. ratio of transmitted RF power to DC power consumed), since they are continuously on. This is generally not acceptable when one is trying to design a handset for a mobile communication system because the usable battery life will be severely reduced. More

complicated linear modulation methods have been devised to allow the use of higher efficiency power amplifiers, but only a few basic techniques are discussed here.

2.1.1.1 Binary Phase Shift Keying

The simplest form of linear modulation is binary phase shift keying (BPSK) where the phase of a constant amplitude carrier is switched between two values (normally 0 and π radians). Then the BPSK signal can be described as

$$\begin{aligned} S_{PSK}(t) &= A \cos(2\pi f_c t) & \text{symbol}=1 \\ S_{PSK}(t) &= A \cos(2\pi f_c t + \pi) & \text{symbol}=0 \end{aligned} \quad (2.1)$$

which can be simplified to

$$S_{PSK}(t) = m(t)A \cos(2\pi f_c t) \quad (2.2)$$

where $m(t)$ is the data signal. If the data is simply a rectangular pulse with amplitude ± 1 , the amplitude of the modulated carrier is constant and only its phase will invert with every change in symbol value as shown in Figure 2.1.

Figure 2.1: Unfiltered PSK modulated carrier.

One might conclude that the BPSK modulated signal is indeed constant amplitude (envelope) and non-linear power amplifiers could be used. However, what is not apparent is that the spectrum of the BPSK signal modulated with rectangular data symbols spreads far beyond the desired channel bandwidth. This is caused by trying to pass rectangular

pulses through a band-limited channel which results in each symbol being spread into adjacent symbol time intervals. This inter-symbol interference (ISI) leads to an increased probability of the receiver making an error in detecting a symbol. Spectral control of the modulated signal is required to simultaneously contain the RF signal within the desired bandwidth and reduce ISI as much as possible. This is accomplished by shaping (filtering) the rectangular data pulses prior to modulation, which makes the modulated carrier depart from constant amplitude. There are many filter types that achieve various degrees of ISI and bandwidth reduction. One of the most popular pulse shaping filters is the raised cosine filter, whose impulse response is shown in Figure 2.2.

Figure 2.2: Raised cosine impulse response.

The spectrum of the unfiltered BPSK signal and the raised cosine filtered signal are shown in Figure 2.3. The null-to-null bandwidth of the unfiltered signal is twice the bit rate with 90% of the energy existing within a bandwidth approximately 1.6 times the symbol rate. The raised cosine filtered signal, on the other hand, contains all of its energy

in a bandwidth 1.5 times the symbol rate. While filtering the baseband signal has contained the spectrum, it does cause the amplitude of the RF signal to fluctuate depending on the filter used.

Figure 2.3: Power spectral density of an unfiltered and raised cosine filtered ($\alpha=0.5$) BPSK signal.

An alternate way of displaying the envelope variation is with a constellation diagram showing the complex envelope of each possible symbol state. The points represent the final magnitude and phase state at the centre of the symbol (i.e. the decision point) while the trajectory connecting them show the path the carrier followed to get from one state to another. A constant envelope modulation scheme would have two or more states connected by a circular path (i.e. constant amplitude). Comparing the constellation diagrams of the rectangular and filtered BPSK signal in Figure 2.4, shows the amplitude variation of the filtered version. This implies that a linear power amplifier must be used to

Figure 2.4: BPSK constellation diagram: (a) unfiltered and (b) filtered.
retain the original BPSK modulated spectrum.

2.1.1.2 Quadrature Phase Shift Keying

Binary phase shift keying (BPSK) modulation is the simplest form of linear modulation, with two defined phase states (0 and π). If the number of states is increased, more data bits per symbol can be transmitted and thus the bandwidth efficiency increases.

Quadrature phase shift keying (QPSK) modulates the carrier using one of four equally spaced phase values (e.g. 0 , $\pi/2$, π and $3\pi/2$) where each value represents a unique pair of data bits. This doubles the bandwidth efficiency of QPSK compared to BPSK. The QPSK signal may be represented by a constant amplitude carrier whose phase is modulated by one of four values

$$S_{QPSK}(t) = A \cos\left(2\pi f_c t + (i-1)\frac{\pi}{2}\right) \quad i = 1, 2, 3, 4 \quad (2.3)$$

Using trigonometric identities, this can be rewritten as

$$\begin{aligned}
 S_{QPSK}(t) = & A \cos \left[(i-1) \frac{\pi}{2} \right] \cos(2\pi f_c t) \\
 & - A \sin \left[(i-1) \frac{\pi}{2} \right] \sin(2\pi f_c t)
 \end{aligned}
 \quad i = 1, 2, 3, 4 \quad (2.4)$$

which expresses the QPSK signal in terms of an in-phase and quadrature (I & Q) component. Based on this representation, a QPSK signal can be depicted using a two dimensional constellation diagram with four points corresponding to the four phase states of the RF carrier as in Figure 2.5.

Figure 2.5: QPSK constellations with different signal sets.

Equation (2.4) exhibits the same constant amplitude characteristics as unfiltered binary phase shift keying. However, as explained earlier, filtering is necessary to contain the signal bandwidth and reduce inter-symbol interference. Baseband filtering of the QPSK signal controls the bandwidth the same way it does for BPSK modulation as shown in Figure 2.6. The consequence of filtering is envelope variation of the modulated signal which is an undesirable side effect. The worst case occurs if the QPSK data sequence causes a π radian phase shift (e.g. $0 \rightarrow 3$) which always occurs in BPSK modulation. Then the carrier amplitude will pass through zero for an instant and any nonlinear

Figure 2.6: Power spectral density of an unfiltered and raised cosine filtered ($\alpha=0.5$) QPSK signal.

amplification will regenerate the filtered side lobes, leading to spectral regrowth. To prevent the regeneration of sidelobes and spectral widening, it is imperative that QPSK (and BPSK) signals be amplified using only linear amplifiers.

There are variants of basic QPSK modulation that eliminate π radian phase shifts, thus preventing the band-limited signal envelope from going to zero (e.g. offset QPSK) but they are still susceptible to some envelope variation. These modified QPSK modulation schemes simply relax the linear amplifier requirement but do not eliminate the problem.

2.1.2 Constant envelope modulation

One of the main concerns with mobile communication systems is power efficiency of the handset. In most cases, the power amplifier efficiency is most important, since it usually determines the amount of transmit time available. The use of non-linear amplifier architectures (e.g. class-C etc.) which have a much higher efficiency than conventional

linear amplifiers is restricted by the type of signal being amplified. The fundamental problem is that linear modulation techniques cannot be used with these amplifiers since they destroy all the baseband filtering used for spectral control and inter-symbol interference reduction. This factor alone has increased efforts to devise new modulation schemes that permit the use of non-linear amplifiers to improve the power efficiency of the handset.

Constant envelope modulation is part of the class of non-linear modulation schemes where the amplitude of the carrier is held constant regardless of the variation in the modulating signal. This has several advantages some of which are:

- Efficient power amplifiers can be used without degrading the spectrum of the transmitted signal.
- Non-coherent discriminator detection can be used which simplifies the receiver design

While constant envelope modulation schemes have many advantages, they occupy a larger bandwidth than linear modulation schemes. In systems where bandwidth efficiency is more important than power efficiency, constant envelope modulation is not well suited.

2.1.2.1 Binary Frequency Shift Keying

In binary frequency shift keying (BFSK), the frequency of a constant amplitude carrier is switched between two values depending on the modulation data. The phase of the transmitted signal may be continuous or discontinuous between bits depending on the way the data is imparted to the carrier. In general, an FSK signal may be represented as

$$\begin{aligned} S_{FSK}(t) &= A \cos(2\pi f_c + 2\pi \Delta f)t & 0 < t < T_b \\ S_{FSK}(t) &= A \cos(2\pi f_c - 2\pi \Delta f)t & 0 < t < T_b \end{aligned} \quad (2.5)$$

for a binary one and zero respectively. The $2\pi\Delta f$ term is a constant offset (deviation) from the nominal carrier frequency. One obvious way to generate an FSK signal is to switch between two oscillators depending on the value of the data. Normally, this form of FSK generation results in a waveform that is discontinuous at the switching points and

therefore is called discontinuous FSK. Phase discontinuities at the switching times pose several problems such as spectral spreading and spurious transmissions, so this type of FSK is generally not used in highly regulated wireless systems [Rapp96]. The more common way to generate FSK is to frequency modulate a single carrier. This is similar to analog FM except that the modulating waveform is now a binary waveform. This type of FSK may be represented as

$$\begin{aligned} S_{FSK}(t) &= A \cos[2\pi f_c t + \phi(t)] \\ &= A \cos\left[2\pi f_c + 2\pi k_f \int_{-\infty}^t m(\tau) d\tau\right] \end{aligned} \quad (2.6)$$

It should be noted that even though the modulating waveform $m(t)$ may be discontinuous at bit transitions, the phase $\phi(t)$ is proportional to the integral of $m(t)$, and is therefore continuous.

The complex envelope of a BFSK signal is a nonlinear function of the data signal so evaluating the spectrum is generally quite involved. The power spectral density of a binary FSK signal consists of discrete frequency components at f_c and $f_c \pm n\Delta f$, where n is an integer. The spectrum of continuous phase FSK falls off as the inverse of the fourth power of the frequency offset from f_c . However, if phase discontinuities exist, the spectrum falls off as the inverse square of the frequency offset from f_c [Couc93]. This makes continuous phase systems more desirable than discontinuous ones. The bandwidth of an FM signal is ideally infinite but Carson's rule gives the approximation

$$B_T \approx 2\Delta f + 2B \quad (2.7)$$

where B is the bandwidth of the digital baseband signal. Assuming that the first null bandwidth is used, a rectangular pulse has a bandwidth equal to the symbol rate R_b . Thus the FSK signal bandwidth for a rectangular pulse becomes

$$B_T = 2(\Delta f + B) \quad (2.8)$$

If raised cosine filtering is used, the FSK signal bandwidth reduces to

$$B_T = 2\Delta f + (1 + \alpha)B \quad (2.9)$$

where α is the attenuation factor of the filter.

2.1.2.2 Minimum Shift Keying

Minimum shift keying (MSK) is a special type of continuous phase frequency shift keying where the modulation index is 0.5. The modulation index is defined as the ratio of the peak frequency deviation to the symbol rate. Thus the peak deviation from the carrier for MSK becomes

$$f_c \pm \Delta f \Rightarrow f_c \pm \frac{1}{4T_b} \quad (2.10)$$

which is one quarter of the symbol rate. The maximum frequency difference between a zero and a one data symbol is exactly one half the symbol rate. This corresponds to the minimum frequency spacing that allows two FSK signals to be coherently orthogonal. The name minimum shift keying refers to the frequency separation required to allow orthogonal detection at the receiver. MSK is spectrally efficient since the frequency spacing used is only half that of conventional non-coherent FSK [Xion94].

MSK modulation belongs to the class of continuous-phase modulation schemes. This implies that the carrier phase does not have any discontinuities that cause the derivative of its phase to be unbounded. This is inherently true, since the frequency of the carrier is modulated rather than its phase and phase is the integral of frequency. Thus any abrupt change in the carrier frequency results in a linear change in its phase over time. During the span of a symbol, the additional change in the carrier phase is called the excess phase and this amount is simply the integral of the frequency deviation over the symbol duration T_b . In MSK modulation, the ± 1 data impulses are zero-order held to produce rectangular pulses. The excess phase due to these pulses is

$$\phi = \int_0^{T_b} f(\tau) d\tau = \pm \frac{\pi}{2T} \int_0^{T_b} 1 d\tau = \pm \frac{\pi}{2} \quad (2.11)$$

This simply states that the excess phase between adjacent symbols is $\pm\pi/2$ radians

depending on the data sequence. Figure 2.7 shows the phase trajectory over time where it clearly shows the change in excess phase due to the symbol sequence. Alternatively, the

Figure 2.7: Phase trajectory of an unfiltered MSK signal.

phase can be easily visualized by viewing the constellation diagram of an MSK signal as in Figure 2.8. Note that due to the constant envelope property of MSK, the phase

Figure 2.8: Constellation of an MSK signal.

trajectory follows a continuous circular path passing through the four phase states. This differs from the discontinuous phase steps of PSK modulation schemes described earlier.

As mentioned earlier, constant envelope modulation schemes have a wider (first null) bandwidth than linear amplitude modulation schemes. This is readily visible in Figure 2.9

Figure 2.9: Power spectral density of MSK and QPSK modulated signals.

where the power spectra of MSK and QPSK signals are compared. What isn't so obvious is that 99% of the MSK power is contained within a bandwidth $1.2/T_b$ while 99% of the QPSK signal power is contained within $8/T_b$ [Rapp96]. This is due to the smoother pulse shapes used, resulting in a faster attenuation of the MSK power. Although the MSK spectrum has lower side-lobe power, the main lobe is wider than the QPSK main lobe so it is spectrally less efficient.

Since there is no abrupt change in phase at bit transition periods, band-limiting the MSK signal does not cause the envelope to go through zero. The envelope is kept more or less constant so efficient nonlinear power amplifiers may be used.

2.1.2.3 Gaussian Minimum Shift Keying

Gaussian minimum shift keying (GMSK) is a derivative of MSK that retains the constant envelope characteristic but also improves the spectral efficiency. This is achieved by shaping the rectangular MSK pulses through additional filtering, which reduces the sidelobes even further. Pre-filtering of the baseband data smooths the phase trajectory of the MSK signal and hence stabilizes the instantaneous frequency variations over time. The consequence of doing this is that the original full response data signal (where each symbol occupies a single bit period) is converted to a partial response signal where each transmitted symbol spans several bit periods. This effectively introduces ISI in the transmitted signal, leading to higher bit error rates and more complicated receiver architectures. GMSK modulation is a compromise between low BER and high spectral efficiency combined in conjunction with constant envelope properties.

The GMSK pre-modulation filter has an impulse response given by

$$h_G(t) = \frac{\sqrt{\pi}}{\alpha} \exp\left(-\frac{\pi^2}{\alpha^2} t^2\right) \quad (2.12)$$

where α is defined by the normalized filter bandwidth BT to be

$$\alpha = \frac{\sqrt{2 \ln(2)}}{BT f_b} \quad (2.13)$$

The filter response can be altered by varying the normalized bandwidth BT . Unfiltered MSK signals are a special case of a Gaussian filter with infinite bandwidth (i.e. $BT = \infty$). Typically, the normalized filter bandwidth is less than one resulting in impulse responses similar to Figure 2.10. Note that as the normalized filter bandwidth BT decreases, the impulse response spreads over adjacent symbols, leading to increased ISI at the receiver. The impact the Gaussian filter has on the spectrum of the MSK signal is visible in Figure 2.11, where the side-lobe power continually decreases with narrower filter bandwidths at the expense of increased ISI.

Gaussian filtering of the baseband MSK data smooths the excess phase trajectory of

Figure 2.10: Impulse response of Gaussian filtered and unfiltered MSK signals.

Figure 2.11: Power spectral density of GMSK signals with various bandwidths.

the carrier. Although the excess phase of a MSK signal contains no discontinuities, it is only piecewise continuous, leading to excessive sidelobes in the MSK spectrum. GMSK signals have a much smoother phase trajectory, but depending on the Gaussian filter bandwidth BT , the excess phase may not reach the desired $\pm\pi/2$ phase shift at the end of a symbol period. This phenomenon is apparent on the phase trellis of Figure 2.12. This

Figure 2.12: Phase trellis of MSK and GMSK signals.

can lead to problems in the receiver since the phase reference over time is lost after a long stream of zeros or ones in the data. Most systems that use GMSK modulation send out a burst of data that restores the phase reference prior to sending the actual data.

The increased ISI in the GMSK signal with narrow filter bandwidth is evident in the constellation diagram of Figure 2.13. The MSK signal constellation in Figure 2.13(a), which has no pre-filtering, has well defined phase states that ease the receiver

Figure 2.13: Constellation of MSK and GMSK signals.

requirements when detecting the symbols. However, the GMSK signal constellation in Figure 2.13(b) shows the phase states are ill-defined and cover a region in each quadrant. This is due to increased ISI from the additional filtering, and the final phase states depend on the Gaussian filter bandwidth and earlier data bit history.

2.2 Constant Envelope Modulators

Low power radio architectures make use of constant-envelope modulation schemes to exploit the use of non-linear power amplifiers which have higher efficiency than linear amplifiers. Constant-envelope modulator architectures have evolved over time mainly due to the continuous thrust toward monolithic integration to reduce size and cost while retaining performance. These modulator architectures may be broadly classified into mixer based, direct modulation and indirect modulation and are described in the following sections.

2.2.1 Mixer Based

Generating a constant envelope MSK signal can be accomplished by combining two quadrature signals that have been appropriately filtered. This is readily apparent by noting that an MSK signal is just a special case of an offset quadrature phase shift keyed signal (OQPSK) with the rectangular pulses replaced by half sinusoid pulses [Pasu79]. An OQPSK signal can be expressed as

$$S_{OQPSK}(t) = a_I(t)\cos(2\pi f_c t) + a_Q(t)\sin(2\pi f_c t) \quad (2.14)$$

where a_I and a_Q are the in-phase and quadrature data streams with values of ± 1 . After half-sinusoidal shaping of the pulses, the RF signal becomes

$$S_{MSK}(t) = a_I(t)\cos\left(\frac{\pi t}{2T_b}\right)\cos(2\pi f_c t) + a_Q(t)\sin\left(\frac{\pi t}{2T_b}\right)\sin(2\pi f_c t) \quad (2.15)$$

Using trigonometric identities, Equation (2.15) can be written as

$$S_{MSK}(t) = \cos\left[2\pi f_c t - a_I(t)a_Q(t)\frac{\pi t}{2T_b} + \phi_k\right] \quad (2.16)$$

where ϕ_k is 0 or π depending on whether a_I is a 1 or -1. From Equation (2.16), a constant

Figure 2.14: Block diagram of a quadrature amplitude modulator (QAM).

envelope is apparent and phase continuity is ensured by choosing a carrier frequency that is an integral multiple of one fourth the bit rate. A modulator structure that realizes this is the quadrature amplitude modulator (QAM) in Figure 2.14 with half-sinusoidal shaped pulses as inputs.

In this architecture, the baseband data is split into an I and Q channel, pulse shaped, and mixed in quadrature with the carrier frequency. Combining both outputs results in the desired MSK signal. Since the data is inherently digital, digital filtering techniques can be used to synthesize the half-sinusoidal shaped pulses before converting them to an analog signal. The drawback to this architecture lies in the up-conversion of the analog baseband I and Q signals. Carrier feed-through from RF mixer offsets and I/Q imbalance (i.e. phase error) in the analog path results in poor sideband suppression.

2.2.2 Direct Modulation

Instead of trying to synthesize a constant envelope signal by combining two amplitude modulated signals (I and Q), one may take advantage of the fact that an MSK signal is simply a frequency modulated (FM) signal. If the data is shaped with the appropriate filter (rectangular for MSK) and followed by an FM modulator, the desired MSK signal results. A relatively simple method of implementing the FM modulator is direct modulation of a voltage controlled oscillator (VCO). The ideal VCO output frequency can be expressed as

$$f_{out} = f_o + K_v V_{tune} \quad (2.17)$$

where f_o is the VCO frequency with 0V input and K_v is the VCO sensitivity in Hz/V. In principle, this will produce an FM signal proportional to the modulating signal. There are several disadvantages in this approach:

- Frequency drift: a change in the VCO frequency caused by tuning voltage drift
- Frequency pushing: a change in the VCO frequency caused by a change in the power supply voltage
- Load pulling: a change in the VCO frequency caused by a change in the VCO load

Many modern digital communication systems use time division multiple access (TDMA) where the transmit time is shared among other users and data is sent in bursts. This implies that the transmitter is inactive for certain time periods. During the inactive time periods, the VCO may be tuned to the desired channel frequency using a phase-locked loop (PLL). When the transmit time occurs, the loop is opened and the VCO can be directly modulated by the data signal. This is known as open-loop modulation and this technique is viable for certain digital communication systems where the transmission time is relatively short. For example, the Digital European Cordless Telephone (DECT) standard has a short ($\sim 500\mu\text{s}$) transmission burst time which prevents excessive VCO

Figure 2.15: DECT open-loop modulator.

frequency drift over time. A typical DECT modulator is shown in Figure 2.15 [Fenk97].

When the transmit function is disabled, the PLL forces the VCO frequency to match the desired channel frequency through closed-loop control. Prior to a transmit burst, modulating data consisting of the DC mean is applied to the VCO input and the loop remains closed to re-lock to the centre frequency. The loop is then opened and the modulation data is transmitted. During the transmit phase, the channel error relies heavily on the stability of the VCO over time and on the residual DC offset supplied by the PLL charge pump after opening the loop. The DECT standard specifies a total frequency error of $<50\text{KHz}$, which is possible for short transmit times only if the PLL charge pump leakage is small. Further complications occur due to secondary effects during the transmit

phase. Switching and power ramping of the power amplifier (PA) cause disturbances of the power supply voltage which indirectly affects the VCO through frequency pushing. A more severe effect is frequency pulling, caused by input impedance changes of the PA when it is switched or ramped [Mohi96]. Reducing the impact of these disturbances on the VCO stability is often a difficult task, thus open-loop modulation is not suitable for standards that have tight frequency control specifications.

2.2.3 Indirect Modulation

The previous section described open-loop modulation of a VCO to generate the MSK signal. Analysis of this technique revealed that it is only a feasible alternative if the transmit times are kept short to prevent VCO frequency drift. There are some digital communication standards that require a longer transmit time and may also demand tighter control over the oscillator frequency and phase (e.g. GSM, DCS-1800 etc.). This precludes the use of open-loop modulation for these applications. However, one desirable feature of open-loop modulation occurs during the non-transmit times, where the PLL loop is closed and the channel frequency is controlled. During this period, none of the previous problems (e.g. frequency drift, load pulling etc.) affecting the VCO frequency persist. If a modulating signal were injected while the PLL loop was locked, accurate channel frequency control could be maintained. Injecting a modulating signal while the PLL loop is closed is known as indirect modulation and various techniques are described in the following sections.

2.2.3.1 Narrow Band Indirect Modulation

Indirect modulation of narrow band signals can be realized by controlling the divider modulus of a $\Delta\Sigma$ fractional-N synthesizer (a type of phase-locked loop) [Rile94]. A typical fractional-N synthesizer is shown in Figure 2.16 and contains a reference oscillator, loop

Figure 2.16: $\Delta\Sigma$ fractional-N synthesizer.

filter, VCO and feedback divider controlled by a $\Delta\Sigma$ modulator. Through closed loop control, the synthesizer output frequency is

$$f_{out} = N f_r \quad (2.18)$$

where N is the divider modulus. For local oscillator applications, the divider modulus N is set to a value that produces a fixed channel frequency. However, if N were made time-varying, the synthesizer would produce a frequency output that follows the instantaneous divider modulus which can be expressed as

$$f_{out}(t) = N(t) f_r \quad (2.19)$$

There are two limitations in this approach; N can only vary by integer increments, resulting in coarse frequency resolution, and the modulating signal bandwidth (BW) is limited. The first problem can be addressed by using a $\Delta\Sigma$ modulator to convert a high resolution, modulating signal into an oversampled low resolution modulus control. An alternative method is to combine $\Delta\Sigma$ techniques into the pulse shaping filter itself, which produces the desired low resolution modulus control as in [Rile94]. The consequence of using either $\Delta\Sigma$ technique to produce high resolution output frequencies from an inherently low resolution synthesizer is the introduction of quantization noise. This noise,

which has been pushed to high frequencies with respect to the reference frequency, must be filtered by adjusting the synthesizer open-loop BW, $\Delta\Sigma$ modulator sample rate and PLL order to set the closed-loop behavior.

The modulation BW limitation can be understood if the synthesizer is viewed as a tracking filter centred at the nominal VCO output frequency with a bandwidth equal to the PLL closed-loop BW seen by the modulating signal. The synthesizer tracks signal frequencies within the closed-loop BW while those outside of the closed-loop BW are suppressed. The advantage of this technique is that no mixers are required to up-convert the modulating signal to the carrier frequency, and the RF signal is inherently band-limited to suppress noise. The downside of this approach is that the modulation BW must be less than the synthesizer BW to avoid any loop suppression of the modulating signal. Since the synthesizer closed-loop BW is usually narrow to suppress the quantization noise of the $\Delta\Sigma$ modulator, the maximum modulation BW is restricted. One method of overcoming this BW limitation is by increasing the synthesizer reference frequency f_r , since the open-loop BW is indirectly governed by f_r . In most cases this is not feasible because this lowers the synthesizer resolution even more, since a unit step of the divider modulus (i.e. the $\Delta\Sigma$ modulator output) now corresponds to a larger frequency step equal to f_r . This can be compensated for by increasing the resolution of the $\Delta\Sigma$ modulator (more complex hardware) to retain the same minimum effective frequency step size through additional dithering. A further penalty is an increase in dynamic power consumption since the $\Delta\Sigma$ modulator is sampled at a higher reference frequency.

2.2.3.2 Offset Phase-Locked Loop

The modulator architecture in [Rile94] generated the GMSK signal by controlling the modulus of a fractional-N synthesizer. Although this produced the desired RF signal for narrow band modulation, it results in a more complex structure due to the need for a $\Delta\Sigma$ modulator. An alternative approach is to fix the divider modulus N to a value corresponding to the desired channel frequency and vary the reference frequency f_r as

Figure 2.17: Modulator with time-varying reference frequency.

shown in Figure 2.17. This implies that the crystal oscillator used as the original reference is replaced by an analog modulator that produces the MSK signal at some intermediate frequency (IF). The synthesizer then functions as a multiplier to shift the modulated IF signal to the desired RF channel frequency. Using Equation (2.19), the synthesizer output becomes

$$f_{out}(t) = N f_r(t) \quad (2.20)$$

where $f_r(t)$ represents the modulated IF signal. Note that the output is now a scaled version of f_r by a factor of N so the frequency deviation at the VCO output has also increased by the same amount. The synthesizer acts like a tracking bandpass filter, which suppresses the noise of the RF output signal. If scaling the modulation frequency deviation in this manner is not acceptable, the divider in Figure 2.17 may be replaced by a mixer and filter. This structure is known as an offset phase-locked loop (OPLL) [Yama97] and depicted in Figure 2.18.

The OPLL operates in closed loop similar to a conventional PLL, except that it compares a frequency offset version of the VCO signal to the reference instead of a divided down version. In doing so, the modulated IF signal is up-converted to the desired

Figure 2.18: Block diagram of an offset phase-locked loop (OPLL).

carrier frequency without altering the modulation. A frequency offset version of the VCO output is generated by mixing the VCO output with an external frequency such that the difference (after being filtered) is equal to the reference (IF) frequency. For example, a GSM (Global System for Communications) transmitter operating in the band 890MHz to 915MHz, can be realized by mixing the RF signal with an oscillator whose range spans 1160MHz to 1185MHz, to yield a 270MHz difference signal. This assumes that the reference signal (modulated by GMSK filtered data) is also 270MHz. Closed-loop control is achieved by comparing the phase of both signals and producing an error signal that depends on the modulating data and the channel frequency.

Several closed-loop bandwidth constraints must be met to ensure adequate performance. A narrow loop bandwidth doesn't provide adequate suppression of the VCO phase noise and may not be sufficient for modulation. However, a wide loop bandwidth, chosen to satisfy the modulation requirements, may result in excessive wideband noise in the transmitted RF signal. This excessive out-of-band noise necessitates additional filtering, commonly implemented using a surface acoustic wave (SAW) filter. Without additional filtering of the modulated RF signal, the wideband noise performance is determined by the OPLL bandwidth. Setting the loop bandwidth to some intermediate value ultimately determines the usable modulating signal bandwidth [Yama97].

The OPLL architecture complexity is similar to the $\Delta\Sigma$ fractional-N modulator except

that the digital $\Delta\Sigma$ modulator and divider is replaced by an analog mixer, filter and IF modulator. There is also the additional requirement for a synthesizer that produces the required offset frequencies for channel selection.

2.2.3.3 Wideband Indirect Modulation

Earlier, it was stated that the fractional-N synthesizer architecture uses indirect modulation to generate the RF signal. Although this technique eliminates the VCO frequency drift inherent in open-loop modulation, the usable modulating signal bandwidth is constrained by the closed-loop bandwidth of the PLL. This is a serious constraint and wideband modulation (with respect to f_r) with adequate quantization noise suppression is generally not possible with this architecture. In [Perr97], a method to compensate for the limited PLL BW of Riley's modulation technique was proposed. The idea was to compensate for the PLL high frequency attenuation by boosting the high frequency components of the modulation signal. After the equalized modulation signal passed through the PLL, the modulation spectrum would be restored to its original form. This wideband architecture is shown in Figure 2.19 and only differs from Figure 2.16 by the inclusion of an embedded digital equalization filter in conjunction with the GMSK pulse shaping filter.

Figure 2.19: Block diagram of a wideband modulator using a digital equalizer and pulse shaping filter.

The noise performance of this architecture is governed by the same criteria as the $\Delta\Sigma$ fractional-N synthesizer in [Rile94]. A key issue is the $\Delta\Sigma$ quantization noise which must be adequately filtered to prevent it from appearing at the VCO output. This implies the PLL loop BW must be sufficiently small and the PLL order high enough to cause the $\Delta\Sigma$ quantization noise to be attenuated at high frequencies. However, the closed-loop BW directly affects the maximum achievable modulation BW (which is the limiting factor in [Rile94]). Therefore it is desirable to control the quantization noise by adjusting the sampling rate of the $\Delta\Sigma$ modulator and choosing an appropriate PLL order. Even if the noise constraints are met, the achievable modulation BW can only exceed the PLL closed-loop BW by prior compensation of the data, as suggested in [Perr97].

The PLL closed-loop transfer function $G(s)$, seen by the modulation data has a low-pass response that limits the overall modulation BW. If the equalizer has a reciprocal frequency response

$$C(s) = \frac{1}{G(s)} \quad (2.21)$$

the modulation transfer function would ideally be flat. The equalizer can be absorbed into the pulse shaping filter to yield one filter with the combined response.

To illustrate the compensation technique described in [Perr97], consider the case where Gaussian frequency shift keying (GFSK) with bandwidth $BT = 0.5$ is used and the PLL is second order. Under these circumstances, the ideal second-order PLL modulation transfer function is

$$G(f) = \frac{1}{1 + \frac{jf}{f_o Q} + \left(\frac{jf}{f_o}\right)^2} \quad (2.22)$$

From Equation (2.22), the equalizer transfer function is

$$C(f) = \frac{1}{G(f)} = 1 + \frac{jf}{f_o Q} + \left(\frac{jf}{f_o}\right)^2 \quad (2.23)$$

The Gaussian filter expressed in the time-domain is given by

$$w(t) = \frac{T}{4} \frac{1}{\sqrt{1.66T_d}} e^{(\pi t / (3.32T_d))^2} \quad (2.24)$$

where T_d is the data symbol period. The equalized filter is obtained by convolving the Gaussian filter impulse response with $c(t)$, the time-domain version of $C(f)$ which gives

$$w_c(t) = w(t) * c(t) = w(t) + \frac{1}{2\pi f_o Q} w'(t) + \left(\frac{1}{2\pi f_o Q} \right)^2 w''(t) \quad (2.25)$$

Substituting for the first and second derivatives of $w(t)$ gives

$$w_c(t) = \left[\left(1 - \frac{1}{1.66 Q f_o T_d} \right) \frac{t}{\sigma} + \frac{1}{(1.66 Q f_o T_d)^2} \left(-1 + \left(\frac{t}{\sigma} \right)^2 \right) \right] w(t) \quad (2.26)$$

where

$$\sigma = \frac{0.833 T_d}{\pi}$$

defines the normalized bandwidth BT of the Gaussian filter.

Equation (2.26) reveals that the signal swing of $w_c(t)$ increases in proportion to $1/(f_o T_d)^2$ for large values of $1/(f_o T_d)$. Since $1/(f_o T_d)$ is the ratio of the modulation data rate and the PLL bandwidth, it is clear that high data rates lead to large signal swings of the modulation signal. If the order of the PLL were increased to n , the signal swing is amplified according to $1/(f_o T_d)^n$ thus compounding the problem.

The dynamic range requirement of the equalized modulation signal imposes the maximum data rate that the PLL can handle. In practice, [Perr97] shows that the phase-frequency detector (PFD) is the limiting component because its dynamic range is limited to one complete reference period, or cycle slipping occurs. Additionally, part of this period is consumed by the dithering action of the $\Delta\Sigma$ modulator, causing the divider phase to bracket that of the reference (i.e. steady-state phase error is never zero).

The attempt to compensate for the analog PLL closed-loop dynamics using digital equalization presents a number of potential problems. The technique used in [Perr97] tries to match the impulse response of the desired continuous-time equalization filter to a

discrete-time version implemented as a finite impulse response (FIR) filter. If enough filter taps are used, the impulse response of the discrete-time filter will not be severely truncated leading to an improper frequency response. However, the magnitude (i.e. tap weights) cannot be matched exactly due to the finite amplitude resolution of the digital filter. This quantization effect may result in a filter that ultimately does not match the desired frequency response.

Assuming a successful digital compensation filter can be realized, it will only match the PLL dynamics for the ideal case. Since the PLL dynamics are analog in nature, they are sensitive to process and temperature variations, leading to mismatch between the compensation filter and PLL closed-loop response. The effect of mismatch is an error in the loop filter poles and zeros and an open-loop gain error due to the unknown VCO sensitivity K_v . Both these phenomena result in an over or under compensated modulation transfer function, leading to inter-symbol interference and frequency deviation error.

Various constant-envelope modulator architectures were introduced in Section 2.2 that are suitable for GMSK modulation. Quadrature amplitude modulators can handle a wide range of modulation schemes but I/Q imbalance and carrier feed-through leads to poor sideband suppression. Direct (open loop) modulation of a VCO is a simple method of constant-envelope modulation but it is difficult to control the carrier frequency so channel drift is a potential problem. For this reason, the technique is restricted to standards that have relatively short transmission times (e.g. DECT). Indirect modulation techniques use closed-loop control of the carrier frequency so drift is not an issue. However, some architectures are more complex (e.g. offset PLL requires separate synthesizer for channel selection), while others do not offer wideband modulation (e.g. $\Delta\Sigma$ fractional-N synthesizer). An equalization technique used to extend the modulation BW of a $\Delta\Sigma$ fractional-N synthesizer was shown to be a viable way of achieving wideband modulation. The consequence of using this technique with analog PLL's is that the PLL characteristics will change due to process and temperature drift. In the next Chapter, a new wideband modulator architecture is introduced that overcomes most of the mismatch problems inherent in the architecture described in [Perr97].

Chapter 3

Continuous-Phase Modulation Using a $\Delta\Sigma$ FD Based Synthesizer

Earlier attempts at wideband modulation of a closed PLL loop [Perr97] revealed several potential problems related to mismatch between the digital compensation filter and the analog PLL characteristics. The inherent process and temperature variations in the analog PLL will shift the open-loop gain and pole/zero frequencies. This introduces mismatch in the form of parasitic pole/zero pairs in the transmit path, causing inter-symbol interference and frequency deviation error.

3.1 Wideband Modulator Architecture

The problem of mismatch between the digital compensation filter and the analog PLL can be minimized (if not eliminated) by using a variant of the PLL architecture first proposed in [Bax95]. The original synthesizer, shown in Figure 3.1, is based on an oversampled $\Delta\Sigma$ frequency discriminator ($\Delta\Sigma$ FD) in the feedback path, which converts the VCO frequency to digital form. The $\Delta\Sigma$ FD functions as a frequency demodulator and analog-to-digital converter, and replaces the divider and phase detector in a conventional fractional-N synthesizer. Early conversion of the VCO frequency into digital form permits much of the remaining loop to be realized using digital signal processing (DSP) techniques. The advantage of this architecture is that most of the loop is digital, yielding predictable and repeatable loop parameters. An added benefit is that analog loop filters are typically off-chip so using a digital loop filter leads to component cost savings.

Operation of this synthesizer architecture is similar to conventional analog PLL's. The

oversampled $\Delta\Sigma$ FD bitstream is decimated and scaled to give a measure of the VCO frequency. This digital value is compared to the desired reference frequency word to produce an error signal. The error signal is digitally filtered and converted to an analog

Figure 3.1: $\Delta\Sigma$ frequency discriminator based synthesizer [Bax95].

signal to drive the VCO. Normally the quantization noise associated with the finite resolution of the digital-to-analog (D/A) converter would introduce noise at the synthesizer output. This problem is overcome by utilizing the noise shaping action of the $\Delta\Sigma$ FD to shift most of the noise energy out-of-band where it is subsequently filtered by the closed-loop response of the synthesizer loop. Although this synthesizer architecture was demonstrated to be feasible for local oscillator applications [Bax94], it is not practical for use as a modulator. This limitation is caused by the reduced sampling rate after decimation, which limits the maximum modulation frequency, and the impracticality of trying to process multi-bit words in real-time. In the following section, a new modulator architecture is proposed that eases the previous restrictions to allow wideband modulation, while simultaneously simplifying the digital signal processing in the loop.

The new modulator architecture [Bax98a] shown in Figure 3.2 is loosely based on the synthesizer described in [Bax95]. It makes use of a $\Delta\Sigma$ FD in the feedback path, but the

low resolution discriminator output (one bit) is processed at the oversampled rate rather than first decimating and filtering its output as in [Bax95]. Retaining the oversampling rate throughout the loop permits a much wider signal bandwidth to be processed before aliasing becomes a problem. The synthesizer makes use of a new modulation technique that directly modulates the $\Delta\Sigma$ FD divider modulus with Gaussian shaped and equalized data. The advantage of this technique is that the output of the $\Delta\Sigma$ FD contains only the error between the VCO frequency and the modulation data, without the DC channel offset. Injecting the modulation data in this way serves two purposes. First, the comparison of the VCO frequency deviation and the modulation data is done within the discriminator, before the quantizer. Thus there is no need to filter and decimate the discriminator output, and subsequently compare it to the high resolution modulation data as in [Bax95]. This also relaxes the dynamic range requirements of the DSP, since the $\Delta\Sigma$ FD output is only one bit wide, whereas in [Bax95] the word size depended on the desired channel resolution. The

power consumption penalty of sampling at the oversampled rate is compensated for by a reduction in complexity of the DSP hardware. A second and more subtle benefit is that $\Delta\Sigma$ modulators (including $\Delta\Sigma$ FD's) may suffer from idle tones caused by DC inputs producing limit cycles. In the case of a $\Delta\Sigma$ FD, a constant input frequency (i.e. zero FM deviation) is analogous to constant amplitude inputs in conventional $\Delta\Sigma$ modulators. Modulating the discriminator modulus input prevents limit cycles from occurring, which would otherwise appear as discrete spurs in the output spectrum. The available dynamic range for direct modulation is limited by the amount of instantaneous phase error the $\Delta\Sigma$ FD can tolerate, since it too is a feedback loop.

Direct modulation of the $\Delta\Sigma$ FD is realized by varying the divider modulus in accordance with the time-varying filtered modulation data. The modulation sampling rate must match the internal sampling rate of the $\Delta\Sigma$ FD, since the divider modulus value is also controlled by the internal feedback path. This modulation technique is viable except that the divider modulus control can only accept a small range of integers (i.e. fixed range of moduli), while the Gaussian filtered and equalized data symbols have a much higher resolution. The conversion of high resolution filtered data to low resolution data can be performed by remodulation of the original signal with a digital $\Delta\Sigma$ modulator. Reducing the resolution inherently produces quantization noise, so the $\Delta\Sigma$ modulator shapes this added noise such that most of its power lies at higher frequencies, and is subsequently filtered by the closed-loop response of the synthesizer.

Normally, the filtered modulation data would be adversely affected by the closed-loop attenuation of the synthesizer, but similar to [Perr97], the data is equalized prior to being injected into the $\Delta\Sigma$ FD. The necessary compensation can be determined by computing the closed-loop transfer function seen by the modulation data and implementing the inverse transfer function. In this architecture, an exact solution can be found since the loop parameters are set in the digital domain and are therefore known. This is a vast improvement over attempting to digitally compensate for an analog PLL which has unknown characteristics due to process and temperature variation as in [Perr97]. Although the poles and zeros of the loop are well defined, the open-loop gain K is not since the sensitivity K_v of the VCO is generally not known. If the open-loop gain error is not

sufficiently reduced, inter-symbol interference and frequency deviation error will exist. A technique to measure and compensate for deviation of the VCO sensitivity from the expected value is described in [Bax98a]. This technique makes use of the existing DSP within the synthesizer loop by digitally correcting for the gain error caused by the VCO. The architecture in [Perr97] does not provide an easy method of correcting any potential gain error although the effects on ISI and frequency deviation error are the same as in this architecture.

3.2 Modelling

Exploration of the modulator design space can be performed by first accurately modelling the behaviour of the various loop components. While models for traditional phase-locked loops are linearized and continuous-time in nature, they are not applicable to the mixed-signal architecture of this synthesizer. Many of the loop components in this architecture are digital and are readily modelled using pure discrete-time models, while the analog blocks can be represented by continuous-time models or optionally be transformed into discrete-time equivalent models. The $\Delta\Sigma$ frequency discriminator will be analysed separately, since it is inherently non-linear and mixed signal in nature with non-uniform sampling.

It is useful to begin with a functional overview of the synthesizer in Figure 3.2 and identify all the relevant signals of interest. When operating as a synthesizer (i.e. modulation input only contains a static channel dithered to provide frequency resolution), the $\Delta\Sigma$ FD produces a noise-shaped frequency error signal. This digital signal is a measure of the VCO deviation from the desired channel frequency, as defined by the difference between the RF and modulation inputs. Note that the discriminator in Figure 3.2 differs from the one in Figure 3.1 due to the additional modulation input. The error signal is digitally processed and subsequently converted to an analog signal that drives the VCO. The synthesizer output frequency is determined by the base modulus of the $\Delta\Sigma$ FD (analogous to the divider modulus in a conventional PLL). In this architecture, the output

frequency is

$$f_{vco} = (n + 1.5)f_r \quad (3.1)$$

where f_r is the reference frequency and n is the $\Delta\Sigma$ FD divider base modulus. The factor of 1.5 is simply an offset caused by the single-loop $\Delta\Sigma$ FD architecture in the synthesizer feedback path, whose operation is explained later in Section 4.1.2. If the modulus n were to vary with time, the $\Delta\Sigma$ FD output would produce a bitstream that represented the instantaneous frequency deviation between the input RF signal and the now time-varying centre frequency $(n(t) + 1.5)f_r$. This relationship assumes that varying the modulus does not cause the $\Delta\Sigma$ FD to lose lock or equivalently, the instantaneous phase error in the $\Delta\Sigma$ FD loop is bounded between $\pm 2\pi$ to prevent cycle slipping. If the modulus n were to track an external signal, the RF carrier would be frequency modulated by that same signal. This is the mechanism used to directly modulate the synthesizer at RF frequencies.

From a functional perspective, this synthesizer operates exactly like a conventional analog fractional-N PLL, but the models must account for the sampled nature of some blocks. In general, the sampling rate (equal to the reference frequency f_r) of the system will be much greater than the modulation signal bandwidth, which simplifies the analysis and reduces the complexity of the models. The central component in this architecture is the $\Delta\Sigma$ FD and a complete understanding of its operation is necessary to further develop an accurate synthesizer model.

3.2.1 $\Delta\Sigma$ Frequency Discriminator Model

The key issue in modelling a $\Delta\Sigma$ FD is the development of an accurate model for the multi-modulus divider. A first-order $\Delta\Sigma$ FD that uses a simple dual-modulus divider will be used to illustrate the principle behind $\Delta\Sigma$ modulation of the frequency rather than the amplitude of a signal. It will be shown that through feedback, the quantization noise introduced by quantizing the phase error is shaped so most of the power lies out of band. The first-order $\Delta\Sigma$ FD architecture will then be extended to the single-loop, second-order structure used in this synthesizer.

Figure 3.3: First-order $\Delta\Sigma$ frequency discriminator (a) model and (b) realization.

A first-order $\Delta\Sigma$ FD, first reported in [Bear94], is composed of a dual-modulus divider, digital phase-frequency detector (PFD) and a quantizer connected in a loop as in Figure 3.3(a). Assuming 1-bit quantization, the PFD and quantizer can be realized with a single 'D' flip-flop clocked by the reference frequency as shown in Figure 3.3(b). The flip-flop acts as a coarse phase quantizer, giving a 1-bit approximation (i.e. the sign) of the phase error. The feedback loop forces the phase of the divider output to track the reference phase in a manner similar to a conventional analog PLL. Due to the 1-bit quantization of the phase error signal, the instantaneous divider output phase brackets the reference phase rather than matching it. This implies that the steady-state phase error never asymptotically reaches zero. The dynamic range of the steady-state phase error cannot exceed the linear range of the phase detector (i.e. $|\phi_e| < 2\pi$) or cycle slipping occurs and the loop loses lock. Although the phase error is being quantized, it can also be expressed as the integrated frequency error since $\phi(t) = \int \omega(t) dt$. Therefore, the discriminator output can

be viewed as an estimate of the instantaneous input frequency deviation sampled at the reference frequency f_r . The analogy of this $\Delta\Sigma$ frequency discriminator to a conventional $\Delta\Sigma$ modulator can be seen by developing an analytical model for a multi-modulus divider (i.e. a dual-modulus divider is the simplest case) and phase detector, using phase instead of amplitude as the variable quantity. The inclusion of a digital PFD complicates the modeling for two reasons:

- it modulates the *width* of its output pulses in proportion to the phase error
- the sampling times are *non-uniform* and depend on whether the phase error is positive or negative for each reference cycle.

These two characteristics are clearly visible in the timing diagram of Figure 3.4 and preclude the use of linear analysis unless some assumptions are made. If the phase error is

Figure 3.4: Digital phase-frequency detector timing diagram.

much smaller than one reference period, an output pulse that precedes the reference edge (due to a negative phase error) will have approximately the same effect as if it was delayed until the reference edge. This in effect, forces the PFD output to be pulses of various width but uniformly sampled at the reference edge, while in practice they are not. Since the PFD now produces an output only at the reference edge, it can be modeled as uniformly sampled impulses whose amplitudes vary in proportion to the phase error (or pulse width) [Craw84]. The fact that the PFD is modeled as a sampled system, provides a convenient coupling mechanism for the divider, since it effectively transports every N^{th} input zero-

crossing to its output. The sampling action of the PFD and the divider can be viewed as a cascade of ideal samplers. Since coherent resampling of an already sampled signal has no additional effect, the divider simply introduces a factor of N^{-1} [Craw94].

The timing diagram in Figure 3.5 illustrates typical waveforms for a first-order $\Delta\Sigma$ FD

Figure 3.5: Timing diagram of first-order $\Delta\Sigma$ frequency discriminator.

using a dual-modulus divider with $n=4$ and a 1-bit quantizer. From Figure 3.5, assuming the divider switches at mid-threshold, the time to the k^{th} divider output corresponding to the j^{th} RF input cycle is

$$t_{d,k} = (T_{vo} + \Delta t_{v,k})j \quad (3.2)$$

where T_{vo} is the nominal RF input period and $\Delta t_{v,k}$ is a time deviation. The j^{th} input cycle occurs at

$$j = nk + \sum_{i=1}^k b_i \quad (3.3)$$

where n is the lower divider modulus and the sum of b_i represents the total number of

extra cycles during which the divider modulus was equal to $(n + 1)$. Thus Equation (3.2) can be expressed as

$$t_{d,k} = (T_{vo} + \Delta t_{v,k}) \cdot \left(nk + \sum_{i=1}^k b_i \right) \quad (3.4)$$

Since phase is the integral of frequency, the divider phase at its k^{th} output with respect to the reference phase is

$$\Phi_{d,k} = \frac{2\pi t_{d,k}}{T_r} \quad (3.5)$$

where T_r is the period of the reference frequency. Substituting Equation (3.4) into (3.5) gives

$$\Phi_{d,k} = \frac{2\pi}{T_r} (T_{vo} + \Delta t_{v,k}) \cdot \left(nk + \sum_{i=1}^k b_i \right) \quad (3.6)$$

Now the average RF frequency (over one reference period T_r) prior to the k^{th} divider output is

$$f_{v,k} = \frac{1}{T_{vo} + \Delta t_{v,k}} \quad (3.7)$$

and defining $N = f_{vo}/f_r$ as the average divider modulus at the centre frequency f_{vo} allows Equation (3.6) to be recast into

$$\Phi_{d,k} = \frac{2\pi}{N} \left[\frac{f_{vo}}{f_{v,k}} \cdot \left(nk + \sum_{i=1}^k b_i \right) \right] \quad (3.8)$$

where $f_{v,k} = f_{vo} + \Delta f_k$ is an angle modulated input signal. Equation (3.8) can be mapped into an equivalent block diagram by recognizing that

$$nk + \sum_{i=1}^k b_i = \sum_{i=1}^k (n + b_i)$$

where the summation is simply an accumulation of RF input cycles over consecutive

divider samples. The f_{vo}/f_v term is an input dependent scaling factor that accounts for the frequency deviation caused by angle modulation of the carrier and any DC channel offset. Minor rearrangement of the terms in Equation (3.8) yields the equivalent block

Figure 3.6: Digital multi-modulus divider model.

diagram of the multi-modulus divider depicted in Figure 3.6.

A complete model of a first-order $\Delta\Sigma$ FD illustrated in Figure 3.7 can now be formed by the inclusion of the sampled PFD, a quantizer and some block manipulation. The divider phase noise and PFD voltage noise are modeled as output referred additive noise sources. Due to the conversion of $f \rightarrow \phi$ in Equation (3.8), there is an ideal integration of frequency into phase sampled at the reference frequency. The quantizer (assuming no overload) is treated as an ideal gain K_q with an additive white noise source. This noise assumption produces good results for higher order ($n > 1$) $\Delta\Sigma$ modulators but fails for simple first-order $\Delta\Sigma$ modulators. Although the quantization noise does not necessarily have white spectral properties, this approach is viable as long as the discriminator does not operate with a constant input (i.e. carrier without modulation). This turns out to be the case in this architecture, since the modulation input n always contains a dithered signal because it's driven by a digital $\Delta\Sigma$ modulator. In the case of single-bit quantization (i.e. two level), the quantizer gain K_q is undefined. One way of circumventing this dilemma is to model the 1-bit quantizer with an effective gain K_q that minimizes the quantization

Figure 3.7: First-order $\Delta\Sigma$ frequency discriminator model.

error power (see Section 4.1.3 for details) [Arda87].

Manipulation of the blocks in Figure 3.7 gives an output equation

$$f_{out} = \frac{N}{f_{vo}} f_v - n + (1 - z^{-1}) S_Q + \frac{N f_v}{2\pi f_{vo}} (1 - z^{-1}) S_{DIV} + \frac{N f_v}{2\pi K_\phi f_{vo}} (1 - z^{-1}) S_{PFD} \quad (3.9)$$

which is composed of the following terms:

- scaled version of the input frequency f_v
- modulation input n
- first-order shaped $\Delta\Sigma$ quantization noise S_Q
- first-order shaped divider phase noise S_{DIV}
- first-order shaped PFD noise S_{PFD}

It becomes clear from Equation (3.9), that the first-order frequency discriminator output bitstream contains signal and $\Delta\Sigma$ quantization noise components resembling that of conventional $\Delta\Sigma$ modulators. This validates its classification as a $\Delta\Sigma$ modulator of

frequency rather than *amplitude*. The ideal differentiation by the $(1 - z^{-1})$ term for the noise sources is due solely to the implied integration from conversion of frequency into phase by the divider. Note that the gain of the divider phase noise and PFD noise is modulated by the input frequency f_v , but this effect is absorbed by the *effective* 1-bit quantizer gain K_q .

Extension of the first-order $\Delta\Sigma$ FD structure into a second-order structure is possible by including one extra integrator in the feed-forward path and a second feed-back path to stabilize the loop. This is illustrated by the classic multi-loop structure in Figure 3.8(a) [Cand85] which has an output

$$out = in + (1 - z^{-1})^2 S_Q$$

giving the desired second-order noise shaping of the quantization noise. Rearranging the

Figure 3.8: Equivalent (a) multi-loop and (b) single-loop second-order $\Delta\Sigma$ modulator structures.

blocks into the single-loop structure in Figure 3.8(b) results in the same signal and noise transfer functions, but with different implications in the hardware implementation.

The first-order $\Delta\Sigma$ FD architecture in Figure 3.7 can be converted into a single-loop, second-order structure similar to Figure 3.8(b) by including an analog integrator after the PFD as shown in Figure 3.9 [Bax96]. The inclusion of the second integrator combined

Figure 3.9: Single-loop second-order $\Delta\Sigma$ frequency discriminator.

with the ideal integration of frequency into phase in the divider, produces the same second-order noise shaping transfer function as the architectures in Figure 3.8. The advantage of the single-loop $\Delta\Sigma$ FD architecture in Figure 3.9 over the multi-loop one described in [Bear94] is that the first integration is acquired for free in the ideal frequency-to-phase conversion in the divider. This eliminates one analog integrator at the expense of a more complicated digital divider (i.e. four-modulus versus dual-modulus) to handle the 2-bit feedback signal generated by the $(2 - z^{-1})$ block.

The second-order discriminator model in Figure 3.10 is formed by adopting the same approach to modeling the first-order $\Delta\Sigma$ FD. It makes use of the same multi-modulus divider model developed earlier and includes the additional integration and $(2 - z^{-1})$ feedback blocks. Note that the second integrator is modeled as a discrete-time block since it is

Figure 3.10: Second-order $\Delta\Sigma$ frequency discriminator model.

driven by the PFD whose output is modeled as a sequence of weighted samples and it's followed by a quantizer sampled at the reference frequency. Similar to the first-order case, the output bitstream can be expressed as

$$\begin{aligned}
 f_{out} = & \frac{N}{f_{vo}} f_v - n + (1 - z^{-1})^2 S_Q + \frac{N f_v}{2\pi f_{vo}} (1 - z^{-1}) S_{DIV} \\
 & + \frac{N f_v}{2\pi K_\phi f_{vo}} (1 - z^{-1}) S_{PFD} + \frac{N f_v}{2\pi K_{CP} K_\phi f_{vo}} (1 - z^{-1})^2 S_{CP}
 \end{aligned} \tag{3.10}$$

where the quantization noise S_Q is now second-order shaped. The divider and PFD noise exhibit the same transfer functions as the first-order model, but now there is an additional noise source due to the analog charge pump. Using the expression for the second-order $\Delta\Sigma$ FD output in Equation (3.10), the complete second-order discriminator model can be simplified to that in Figure 3.11. The purpose of this simplified model is to replace the

Figure 3.11: Linear equivalent noise model of a second-order $\Delta\Sigma$ frequency discriminator.

non-linear operation of the $\Delta\Sigma$ FD loop with a linear open-loop model that retains the same functionality and dominant noise sources. The linear model in Figure 3.11 can now be used to represent the $\Delta\Sigma$ FD in the main synthesizer model described next.

3.2.2 Synthesizer Model

Having developed a simplified linear model for the $\Delta\Sigma$ FD, the rest of the synthesizer modelling can be completed. As mentioned earlier, the digital signal processing is easily represented by discrete-time Z-domain models. Without knowing the exact implementation of the DSP, errors due to round-off, overflow, etc., cannot be modelled, so an ideal discrete-time model (i.e. infinite precision and range) will be used. The remaining synthesizer loop components; D/A converter, charge pump and VCO can be represented by their corresponding continuous-time S-domain models. The interface between the digital and analog blocks reduces to a conversion gain followed by a zero-order hold function in the D/A converter. This provides the necessary conversion of weighted discrete-time samples to pulses of finite duration and amplitude.

The $\Delta\Sigma$ FD, modeled as a discrete-time linear system, samples the VCO frequency and produces a quantized sampled output. However, sampling the continuous-time VCO output will replicate copies of the VCO noise spectrum about integer multiples of the

sampling rate (in this case the reference frequency f_r). That is to say, if $f(t)$ is a signal whose Laplace transform is $F(s)$, its sampled equivalent is

$$f^*(t) = \sum_{n=-\infty}^{\infty} f(t)\delta(t - nT) \quad (3.11)$$

where $\delta(t)$ is the dirac delta function. Taking the Laplace transform of Equation (3.11) gives

$$F^*(s) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(s - jn\omega_s) \quad (3.12)$$

which contains multiple copies of the original signal spectrum $F(s)$ at integer multiples of ω_s and scaled by $1/T$. If the phase noise spectrum of the VCO is band limited to $f_c \pm f_r/2$, or at least attenuated at frequency offsets $f > f_r$, the amount of aliasing due to sampling will be negligible. This assumption is true if the synthesizer open-loop bandwidth is much less than the reference frequency (i.e. $BW < 0.1 f_r$) which is generally the case. The complete synthesizer model, shown in Figure 3.12, combines all the main

Figure 3.12: Linearized equivalent model of $\Delta\Sigma$ FD based synthesizer.

loop component models except those in the modulation path which are not part of the loop.

The modulation path, shown in Figure 3.13, consists of a digital transmit filter that implements the desired Gaussian pulse shaping and equalization of the synthesizer closed-loop response, and a digital $\Delta\Sigma$ modulator. Pre-shaping of the data symbols is necessary to

Figure 3.13: Modulation data path.

control the RF output spectrum of the synthesizer when modulated. In digital radio applications, a common standard, as discussed in Chapter 2, is Gaussian minimum shift keying (GMSK), which provides smooth transitions (continuous phase) between symbol values through Gaussian filtering. The consequence of pre-shaping is that it introduces inter-symbol interference (ISI) that causes spreading of adjacent symbols so they overlap in time. This implies that there is a compromise between the desired RF signal bandwidth and the amount of ISI the receiver can tolerate while maintaining a desired bit error rate (BER). From a modeling perspective, a general Gaussian filter can be described which has a normalized bandwidth BT determined by the modulation standard used. This assumes that the input data symbols are finite width pulses of duration T_b and sampled at the reference frequency f_r . Since the bipolar data symbols can be uniquely described by impulses (i.e. ± 1), it is convenient to absorb the rectangular pulse, denoted by $r(t)$, into the Gaussian filter. In the time domain, this implies convolving the two responses to give

$$\begin{aligned}
 g(t) &= r(t) * h(t) \\
 &= [u(t) - u(t - T_b)] * \left[\frac{1}{\sqrt{\pi}} \beta e^{-[(t + T_b/2)\beta]^2} \right]
 \end{aligned} \tag{3.13}$$

where β is defined as

$$\beta = \sqrt{\frac{2}{\ln(2)}} \pi f_b B T$$

Equation (3.13) can be recast in terms of error functions so that

$$g(t) = \frac{1}{2} \{ \text{erf}[\beta(t + T_b/2)] - \text{erf}[t - T_b/2] \} \quad (3.14)$$

Implementing Equation (3.14) as a digital finite impulse response (FIR) filter requires the Gaussian impulse response $g(t)$ to be sampled at a rate compatible with the rest of the system, which in this case is the reference frequency f_r . The required FIR filter tap weights are extracted directly from the sampled impulse response $g(nT_b)$ and quantized to a fixed number of bits. The resulting transfer function can be modeled in the Z-domain in terms of the quantized tap weights b_n as

$$H_{GAUSS}(z) = b_1 + b_2 z^{-1} + b_3 z^{-2} + \dots + b_n z^{-(n-1)} \quad (3.15)$$

where the filter order is chosen to minimize any truncation effects.

The second portion of the modulation filter compensates for any in-band ripple (e.g. second order peaking) and out-of-band attenuation of the synthesizer closed-loop response. The need for equalization is a direct consequence of the modulation signal bandwidth exceeding the synthesizer closed-loop bandwidth. The first step in developing an equalizer model is to compute the output transfer function seen by the modulating signal which controls the $\Delta\Sigma$ FD modulus n . This is readily accomplished by manipulating the blocks of the previously developed synthesizer model in Figure 3.12 to get

$$H_{CHAN}(z, s) = \frac{H_{VCO}(s) \cdot H_{CP}(s) \cdot H_{DAC}(s) \cdot H_{DSP}(z)}{1 + G(z, s)} \quad (3.16)$$

where $G(z, s)$ is the synthesizer open-loop response. Note that the transfer function is defined using the S-domain and Z-domain due to the mixed-mode signals present in the synthesizer. The modulation closed-loop magnitude response is computed by solving for

$$|H_{CHAN}(z, s)|_{s=j\omega, z=e^{j\omega T_r}}$$

where T_r is the sampling period. Compensation for the effect of $H_{CHAN}(z, s)$ in the modulation path is accomplished by cascading an equalizer after the Gaussian filter with a reciprocal response

$$H_{EQ}(z) = \frac{1}{H_{CHAN}(z, s)} \quad (3.17)$$

The concern over digital compensation of a mixed signal transfer function that arose in [Perr97] is not an issue here because all of the synthesizer blocks are digital except for the analog charge pump and VCO. Solving for $H_{EQ}(z)$ in Equation (3.17) is non-trivial because of the mixed signal representation of the blocks. One may choose to map all the continuous-time blocks into their equivalent discrete-time blocks (aliasing due to sampling forbids mapping $Z \rightarrow S$). The choice remains whether to match the impulse or frequency response of the required equalizer transfer function, since aliasing in sampled systems often prevents matching both responses simultaneously [HP243]. In the context of a wideband modulator, the frequency response is most important, so this approach will be used here. Analytical mapping from the S-domain to the Z-domain can be achieved using a bi-linear transform which provides good matching at low frequencies (i.e. $f \ll f_r$) but contains significant distortion (i.e. warping) at higher frequencies. Rather than attempting an analytical solution, it is easier to utilize one of the many filter synthesis functions that are made available in filter design software packages. These synthesis functions try to minimize the error between the synthesized filter response and the desired response according to some cost function. This approach is attractive in the sense that often an exact match is not necessary for frequencies outside of a certain bandwidth. Thus the synthesis tool can apply more weight to minimize the error in the bandwidth of interest at the expense of larger out-of-band error. This has the advantage of yielding a filter of lower order, which is easier and less costly to implement. Realization of the digital filters can be simplified by combining the Gaussian and equalizer responses, since their opposing transfer functions will result in a smaller required dynamic range. A further improvement is possible by synthesizing a finite impulse response (FIR) filter for both transfer functions, which allows read-only-memory (ROM) based filter architectures to be used.

The advantage of the ROM based filter architecture is explained in the modulator implementation of Chapter 5.

3.3 Design Parameters

Previously, models of all the synthesizer blocks were developed based on the proposed architecture for a wideband modulator. What remains to be determined is the value of the various loop parameters that achieve the desired modulation, while meeting the synthesizer noise and transient specifications. GSM modulation is chosen as a design example although the architecture is applicable to any wideband, continuous-phase, constant envelope modulation scheme. The GSM modulation standard is a constant envelope GMSK modulation with a symbol rate of 270.833 Kb/s, normalized bandwidth $BT = 0.3$, and a maximum frequency deviation of ± 67.71 KHz.

As mentioned in [Perr97], the modulation and noise bandwidth requirements may be decoupled by using an equalizer in the modulation path. This is an important feature, since noise and modulation specifications usually result in conflicting synthesizer loop parameters (i.e modulation $\Rightarrow$ wide loop BW, noise $\Rightarrow$ narrow loop BW). Taking advantage of this extra degree of freedom, the first step is to compute the loop parameters that are required to meet the noise specification. The goal is to realize the simplest architecture that performs to the desired specification, leading to lower cost and overall power consumption.

3.3.1 PLL Order

While there are many reasons for choosing a particular PLL order, conventional fractional-N architectures impose a *minimum* PLL order necessary to obtain sufficient filtering of the out-of-band quantization noise. Conventional $\Delta\Sigma$ fractional-N synthesizers require a PLL order equal or greater than the $\Delta\Sigma$ modulator order to attenuate the quantization noise outside the loop BW. If the PLL order is equal to the $\Delta\Sigma$ modulator order, the inherent

integration of frequency into phase gives a net phase noise attenuation of -20dB/dec outside the loop BW.

In this architecture, a second-order $\Delta\Sigma$ FD is chosen to provide sufficient noise shaping without resorting to higher order structures that are inherently more difficult to implement [Rile98]. Unlike conventional $\Delta\Sigma$ fractional-N synthesizers, the quantization noise in this architecture is subjected to a different noise transfer function (see Equation (5.8)) to the output due to the $\Delta\Sigma$ FD in the feedback path. Assuming a second-order synthesizer (loop filter is an integrator with phase lead compensation), the noise transfer function attenuates within the loop BW at -20dB/dec. Since the *frequency* quantization noise for a second-order $\Delta\Sigma$ FD rises at 40dB/dec, the equivalent phase noise rises at 20dB/dec similar to conventional $\Delta\Sigma$ synthesizers of equal order. While this may seem problematic, the PLL loop BW may be set sufficiently low (in both architectures) so the total quantization noise integrated over the band of interest is attenuated due to the in-band $\Delta\Sigma$ noise shaping and the closed-loop suppression out-of-band. Additional quantization noise suppression is possible by increasing the order of the synthesizer, which complicates the stability for orders greater than two, or by providing additional filtering outside the loop BW. Rather than increasing the synthesizer order to three, a second-order synthesizer with additional quantization noise filtering will be used. Note that the digital $\Delta\Sigma$ modulator in the modulation path is subjected to the same transfer function as the $\Delta\Sigma$ FD (this is shown in Section 5.4), so its order must be no greater than that of the $\Delta\Sigma$ FD. The combined quantization noise from the $\Delta\Sigma$ modulator and the $\Delta\Sigma$ FD are filtered by the same PLL closed-loop transfer function.

3.3.2 Reference Frequency

The choice of reference frequency in this synthesizer influences a number of parameters. Most importantly, it defines the sampling rate for the $\Delta\Sigma$ FD and all the DSP following it. The sampling rate with respect to the loop BW must be sufficiently high to prevent the effects of aliasing (note: any synthesizer with a digital divider in the feedback path is affected by aliasing). Aliasing in discrete-time PLL's increases the open-loop gain over

that in equivalent continuous-time systems. This is most pronounced at one half the sampling frequency $f_s/2$, where the baseband and replicated transfer function overlap (i.e. alias) thereby reducing the available phase margin [Craw84]. Restricting the loop BW to be less than $0.1 f_s$ reduces the effects of aliasing to negligible levels, so this imposes a minimum reference frequency for a given loop BW.

Further constraints on the minimum reference frequency are due to the modulation data path. Recall that the modulation BW, by definition, is greater than the loop BW, so equalization is necessary to restore the modulation signal after passing through the synthesizer. The amount of allowable equalization is determined by the input dynamic range of the $\Delta\Sigma$ FD, since the modulation is injected into the discriminator (refer to Figure 3.2). This limit, defined in Chapter 4, is equal to the sampling frequency f_s . Without equalization, the required modulation frequency dynamic range is equal to the frequency deviation of the adopted standard (e.g. ± 67.71 KHz for GSM). However, depending on the out-of-band attenuation of the synthesizer, a large amount of equalization may be necessary for higher signal frequencies, which quickly consumes the available dynamic range. A second modulation constraint arises from the use of $\Delta\Sigma$ noise shaping to reduce the high resolution of the GMSK filtered data symbols. As with any noise-shaped encoding, the usable BW (to maintain a desired SNR) is defined to be a fraction of the sampling frequency or equivalently, a minimum oversampling ratio (OSR).

3.3.3 Stability

Having defined the overall order of the PLL and $\Delta\Sigma$ FD, the loop dynamics can now be tailored to yield the desired stability and transient response. In striving for maximum modulation dynamic range, it is important to ensure that the synthesizer operates with a near zero steady-state phase error. Any residual phase error will encroach on the available dynamic range and limit the amount of equalization the synthesizer can tolerate (i.e. modulation BW is reduced). Note that in this synthesizer, varying the channel frequency has no effect on the loop dynamics since the $\Delta\Sigma$ FD only produces an error while excluding any DC offsets. This does not hold in conventional $\Delta\Sigma$ fractional-N synthesizers where

large changes in channel frequency (i.e. controlled by divider modulus n) will alter the dynamics of the PLL. Reducing the residual phase error amounts to having a high DC gain in the forward path, which can be realized with an integrator.

Many methods exist to evaluate the stability of a linear time-invariant system, including the Routh-Hurwitz criterion, the Nyquist criterion, root-locus methods and Bode diagrams [Kuo75]. The synthesizer models developed earlier are mixed-signal in nature and present a few complications to resolve before the stability analysis can begin. Continuous S-domain models may be mapped into their equivalent discrete-time representation and the resulting Z-domain system model can be analyzed. This tends to be tedious and somewhat inaccurate because only the frequency *or* impulse response can generally be matched. Alternatively, modeling the loop stability of this mixed-mode synthesizer using pseudo-continuous modeling techniques provides a close approximation to the true open-loop gain provided the system BW is a small fraction of the sampling frequency. This was the approach used in model development, where the loop BW was assumed to be much less than the reference frequency. The stability of the synthesizer may then be determined by evaluating the open-loop response while optionally including the effect of the replicated copies due to sampling. Ignoring sampling effects will still provide a reasonably accurate measure of loop stability (i.e. actual phase margin will be less than predicted) as long as the open-loop bandwidth $BW \ll f_r$.

Stability analysis is accomplished by examining the synthesizer open-loop gain through the use of Bode plots in conjunction with the pseudo-continuous models developed earlier. A high gain in the synthesizer forward path is realized by using an integrator with phase lead compensation as a loop filter. The compensating zero is necessary to ensure an adequate phase margin at the open-loop unity gain frequency. Additional filtering outside of the loop BW provides more quantization noise attenuation with minimal effect on the loop stability. Since the complete loop filter is realized in the digital domain, it is desirable to find equivalent equations that relate the desired synthesizer parameters (e.g. BW, damping factor etc.) to the actual circuit parameters (e.g. discrete-time filter gain, poles, zeros etc.).

A continuous-time integrator with phase lead compensation can be expressed as

$$F(s) = \frac{1 + \tau_2 s}{\tau_1 s} \quad (3.18)$$

which has a zero at $\omega_z = 1/\tau_2$ and a high frequency gain (i.e. as $s \rightarrow \infty$) equal to τ_2/τ_1 [Blan92].

A second-order continuous-time PLL has an open-loop transfer function $G(s)$

$$G(s) = \frac{K'(1 + \tau_2 s)}{\tau_1 s^2} \quad (3.19)$$

where K' is the open-loop gain and the extra integration term arises from the frequency to phase conversion in the VCO. A plot of $G(s)$ in Figure 3.14 illustrates the effect the

Figure 3.14: Second-order synthesizer open-loop Bode plot.

compensating zero has by reducing the gain slope and phase as the gain magnitude passes the 0dB point. It is common to express a second-order PLL in terms of its natural

frequency ω_n and damping factor ζ . However, when designing an actual synthesizer, the closed-loop bandwidth and damping factor are more useful parameters. From Figure 3.14, the natural frequency and damping factor in terms of the loop filter time constants are

$$\omega_n = \sqrt{\frac{K'}{\tau_1}} \quad (3.20)$$

$$\zeta = \frac{\tau_2}{2} \sqrt{\frac{K'}{\tau_1}} \quad (3.21)$$

The synthesizer closed-loop BW is approximately equal to the open-loop unity gain frequency, which in this case is

$$\text{BW} = K' \frac{\tau_2}{\tau_1} \quad [\text{rad/s}] \quad (3.22)$$

Although this mixed signal synthesizer behaves differently from its continuous-time counterpart due to sampling, at frequencies much less than f_r , its behavior is approximately the same, so one may make use of the same loop parameters.

The discrete-time equivalent filter is composed of a proportional term a_2 and an integral term a_3 with a transfer function

$$F(z) = a_2 + \frac{a_3}{1 - z^{-1}} \quad (3.23)$$

This can be rearranged into standard form to give

$$F(z) = (a_2 + a_3) \cdot \left(\frac{z - \frac{a_2}{a_2 + a_3}}{z - 1} \right) \quad (3.24)$$

Equation (3.24) shows that the digital filter has a zero at $z = a_2/(a_2 + a_3)$ and a pole at $z = 1$. In the frequency domain, this is equivalent to an integrator with a zero at

$$\omega_z = -f_r \ln \left(\frac{a_2}{a_2 + a_3} \right) \quad [\text{rad/s}]$$

and a high frequency gain of $(a_2 + a_3)$.

Computing the mixed-signal equivalent natural frequency ω_n is possible by using Equation (3.20) and finding an equivalent expression for the loop filter time constant τ_1 . Using the continuous-time and discrete-time filter parameters, τ_1 becomes

$$\begin{aligned}\tau_1 &= \frac{\tau_1}{\tau_2} \tau_2 \\ &= \frac{1}{(a_2 + a_3)} \cdot \frac{1}{-f_r \ln\left(\frac{a_2}{a_2 + a_3}\right)}\end{aligned}$$

The natural frequency ω_n for the discrete-time case is

$$\begin{aligned}\omega_n &= \sqrt{\frac{K'}{\tau_1}} \\ &= \sqrt{-K' f_r \ln\left[\left(\frac{a_2}{a_2 + a_3}\right)^{(a_2 + a_3)}\right]} \quad [\text{rad/s}]\end{aligned}$$

In a similar manner, the equivalent discrete-time damping factor is found to be

$$\begin{aligned}\zeta &= \frac{\tau_2}{2} \sqrt{\frac{K'}{\tau_1}} \\ &= \frac{\sqrt{K'(a_2 + a_3)}}{2 \sqrt{-f_r \ln\left(\frac{a_2}{a_2 + a_3}\right)}}\end{aligned}$$

while the equivalent closed-loop BW is

$$\begin{aligned}\text{BW} &= K' \frac{\tau_2}{\tau_1} \quad [\text{rad/s}] \\ &= K'(a_2 + a_3)\end{aligned}$$

Due to the discrete-time nature of many synthesizer blocks, the open-loop phase will rapidly change for frequencies approaching the sampling frequency, unlike in the equivalent continuous-time synthesizer. This phenomenon forces one to use the equivalent discrete-time damping factor ζ with caution. A more meaningful measure of the loop

stability is the phase margin, which in this case is difficult to compute analytically but relatively easy to evaluate numerically.

Additional attenuation of the $\Delta\Sigma$ FD quantization noise is achieved by introducing a lowpass filter whose poles are set outside the loop BW. Since this filter is also realized in DSP, it will further erode the phase margin, but its effect can be controlled by careful setting of the synthesizer loop parameters. A typical open-loop transfer function is plotted in Figure 3.15 where the effect of sampling is clearly visible by the decrease in phase

Figure 3.15: Comparison between discrete-time (13MHz sampling frequency) and continuous-time synthesizer open-loop transfer functions.

margin outside of the loop BW. The equivalent continuous-time PLL would have a near 90 degree phase margin due to the compensating zero, limited only by the extra filter poles providing additional out-of-band noise attenuation.

Defining the loop parameters begins with satisfying the phase noise requirements, which implies setting the loop BW once the PLL order is determined. Once the loop BW

is set, the stability can be evaluated and if necessary, adjusted to yield an acceptable phase margin. In Chapter 5, a modulator design for a particular application will be developed and actual loop parameter values will be computed.

3.4 Equalization of Synthesizer Closed-Loop Response

High data rate modulation of a synthesizer with a much narrower loop BW is made possible by compensating for the inherent attenuation of the closed-loop response. This allows the loop BW and modulation BW to be independently set to achieve a desired noise performance for a given data rate.

3.4.1 Transmit Filter

The filtering requirements can be decomposed into two functions. First pre-filtering of the digital data symbols is necessary to control the spectrum of the RF signal. This requirement is usually defined by the modulation scheme chosen and involves band limiting the symbol pulses through a filter, which in this application has a Gaussian response. Band limiting the baseband data implies using a filter with a lowpass response. The second requirement is to compensate for the attenuation of the closed-loop transfer function of the modulation path. This form of compensation yields a filter with a highpass response (i.e the inverse of the closed-loop transfer function). Considering each filter separately, one finds that the Gaussian filter (depending on the normalized bandwidth BT) may have stringent out-of-band attenuation requirements, while the equalizer filter has a high gain for frequencies outside of the synthesizer loop BW. When realizing these filters independently, the dynamic range requirements will have to be large to handle the wide range of signal amplitudes, which leads to complex filter structures. Since both filters are cascaded, it is advantageous to combine the responses of both filters into one as shown in

Figure 3.16: Transmit filter composed of Gaussian and equalizer responses.

Figure 3.16. In doing so, the dynamic range of the filter is reduced since the Gaussian filter attenuation neutralizes the equalizer high frequency amplification which results in a simpler implementation. In cases where the modulation closed-loop transfer function is heavily attenuated, combining the Gaussian and equalizer filter responses may still not result in a small enough dynamic range due to the excessive high frequency compensation required. This situation arises since the equalizer transfer function is compelled to

Figure 3.17: Baseband filter response with (a) unrestricted and (b) restricted modulation bandwidths.

compensate for the closed-loop response over the entire bandwidth $0 < f < f_r/2$ to guarantee a flat modulation channel response as shown in Figure 3.17(a). In most cases,

one can tolerate some band limiting and still achieve the desired spectral containment of the modulated RF signal. Using this approach, an equalization bandwidth less than $f_r/2$ can be defined that allows the combined response of the filters to be more uniform at higher frequencies, as shown in Figure 3.17(b).

The type of filter used to realize the Gaussian and equalization filter responses largely depends on the desired final implementation. Choosing an infinite impulse response (IIR) filter usually realizes the desired response with a lower filter order than the equivalent finite impulse response (FIR) filter. The drawback of an IIR filter is that it must be realized in hardware using arithmetic blocks operating at the sampling frequency. Alternatively, a corresponding FIR filter will have a much longer impulse response (i.e. more taps) but it is possible to realize it in a ROM architecture, which reduces complicated computations to a simple table lookup of pre-computed values. The implication of choosing a FIR filter is that the resulting filter must be constrained to having only zeros. As discussed during the model development, the easiest method to synthesize the filter is to use a filter synthesis algorithm rather than trying to directly find the impulse response of the combined Gaussian and equalizer filter transfer functions. This method allows the equalization bandwidth to receive a higher weighting than the out-of-band frequencies, resulting in less error where the effect of the equalizer is more sensitive on the overall output spectrum.

3.4.2 Dynamic Range Constraints

Without applying modulation, the minimum dynamic range of the synthesizer reduces to that required to process the out-of-band noise from the $\Delta\Sigma$ FD in addition to providing the basic tuning range for channel selection. This does not include the extra range necessary when switching channels since this can be handled by special acquisition modes that optimize the switching speed through loop parameter variation. Consider the synthesizer block diagram in Figure 3.18. It is convenient to separate the $\Delta\Sigma$ FD from the rest of the blocks because the $\Delta\Sigma$ FD is essentially a control loop within the larger synthesizer loop.

Figure 3.18: Block diagram of the $\Delta\Sigma$ FD based GMSK modulator.

Since the $\Delta\Sigma$ FD produces a single-bit output representing the frequency deviation error between the RF and modulation inputs, the dynamic range requirement of the DSP is solely defined by the synthesizer loop parameters that dictate the necessary filter transfer function. Since the filters are entirely digital, the dynamic range is set by the word size of each internal node, which can be arbitrarily increased. The obvious penalty of processing a large word size is that the complexity of the DSP increases dramatically, leading to increased area requirements. Even if the internal dynamic range requirements of the DSP are met through scaling, the output word may still be large, which forces the digital-to-analog (D/A) converter range to be at least as large. Designing a high resolution D/A converter to handle the dynamic range is generally not trivial since it is difficult to restrain the differential and integral linearity to acceptable limits. If the dynamic range scaling efforts were applied to both the digital *and* analog blocks, there may be a viable solution where both the digital word size and the analog charge pump conversion gain are acceptable. If that approach is not feasible (i.e. cannot be realized effectively in hardware), an alternative method is to constrain the dynamic range of the DSP output, thereby relaxing the D/A converter requirements. Reducing the dynamic range can be

accomplished by remodulating some of the least significant bits (LSB's) with a $\Delta\Sigma$ modulator as illustrated in Figure 3.19. This technique effectively quantizes some of the

Figure 3.19: Reducing the D/A dynamic range through remodulation.

high resolution LSB's into a lower resolution bitstream that dithers the next higher significant bit (with potential carry overflow). This introduces a truncation error but the resulting noise is shaped by the $\Delta\Sigma$ modulator and the amount of quantization noise is much less than that produced by the $\Delta\Sigma$ FD. For example, if a 16-bit word must be reduced to a 12-bit word, the $\Delta\Sigma$ modulator must remodulate the lower four LSB's and its output must drive the fifth LSB. The amount of additional quantization noise introduced depends on the quantization level Δ used, which in this case is $2^4 = 16$. This corresponds to a total noise power

$$S_{Q(RMS)}^2 = \frac{\Delta^2}{12} = \frac{(2^4)^2}{12} \quad (3.25)$$

which is $(2^{16-4})^2 = (4096)^2$ times less than if the entire word were quantized into 1-bit.

3.4.2.1 $\Delta\Sigma$ Frequency Discriminator Overload

Dynamic range issues pertaining to the $\Delta\Sigma$ FD are classified for the two modes of operation — when the synthesizer is used as a local oscillator and when used as a transmit modulator. In the synthesizer mode, the RF input to the $\Delta\Sigma$ FD during steady-state conditions varies only due to the phase noise of the VCO, which is assumed to have a

small effect on the carrier frequency. If the carrier frequency lies within the $\Delta\Sigma$ FD input dynamic range, as described in Section 4.1.2 (i.e. $(n + 1)f_r < f < (n + 2)f_r$), the output bitstream will have an average value

$$\overline{f_{out}(t)} = 0.5 + \frac{1}{f_r} [f_{in}(t) - (n + 1.5)f_r] \quad (3.26)$$

Since the input frequency range is f_r , it is intuitive that the carrier frequency fluctuations due to phase noise are much less, and therefore pose no dynamic range problem.

If modulation is applied directly to the $\Delta\Sigma$ FD modulus control, the dynamic range requirements are not as obvious compared to the simple synthesizer mode. The problem becomes clearer if one recognizes that the $\Delta\Sigma$ FD actually has *two* signal inputs — the RF carrier and the modulation control, and the discriminator responds to the relative difference between them. In the previous case, the modulus control n was fixed to represent a channel frequency while the RF carrier was fluctuating due to phase noise. However, as a modulator, the modulus control is allowed to vary according to the filtered modulation data, so the $\Delta\Sigma$ FD output is now defined as

$$\overline{f_{out}(t)} = 0.5 + \frac{1}{f_r} [f_{in}(t) - (n(t) + 1.5)f_r] \quad (3.27)$$

where $n(t)$ is now time varying. One may be perceptive and argue that due to closed-loop control, the VCO frequency will follow that of the up-banded filtered modulation data and the net difference seen by the discriminator would be identical to the synthesizer case with constant RF and modulus inputs. This would indeed be true if the synthesizer could track the actual modulation data, but due to its limited loop BW, it cannot do so without compensation. The need for compensation increases the dynamic range of the filtered modulation data, as seen in Chapter 2, but the VCO still follows the original unmodified modulation data. Now the modulus control n and the RF input do not track each other and this appears to the discriminator as a signal with a larger *relative* dynamic range. While the frequency deviation of the RF carrier due to modulation is determined by the chosen modulation scheme, the dynamic range of the equalized modulation data is not readily known since it depends on the loop parameters. If the loop parameters are set to satisfy a

desired modulation scheme, then an appropriate Gaussian and equalizer filter can be synthesized. Once the filter is known, a random data sequence is convolved with the filter's impulse response, and the peak-to-peak amplitude of the output can be determined. This is difficult to do analytically due to the random nature of the sampled data sequence B_n , but numerically it is trivial to compute:

$$\text{dynamic range} = \left\| \sum_{n=-\infty}^{\infty} B_n(t - nT_r) * h_{EQ}(nT_r) \right\| \quad (3.28)$$

The available dynamic range of the $\Delta\Sigma$ FD is actually less than f_r , since constant frequencies at the two extremes of the $\Delta\Sigma$ FD input range would produce an error sequence of all zeros or ones. A stream of bits with constant value obviously has no noise shaping, so in practice the usable dynamic range is somewhat less (typically 80% of the theoretical maximum to prevent quantizer overload). Note that temporary excursions of the input frequency (e.g. a few concurrent samples) that exceed the dynamic range are tolerated as long as the $\Delta\Sigma$ FD can keep the error bounded so its quantizer doesn't saturate.

3.4.2.2 Digital $\Delta\Sigma$ Modulator Range

Coupling of the modulation data into the synthesizer is realized by controlling the divider modulus of the $\Delta\Sigma$ FD. The previous section defined the discriminator input dynamic range to be equal to the reference frequency f_r , which corresponds to changing the divider modulus from n to $(n + 1)$ where n is an integer. The mechanism for converting the high resolution Gaussian filtered data into low resolution integers is by $\Delta\Sigma$ modulation. One may view the $\Delta\Sigma$ modulator as a necessary but undesirable block in the modulation path since it only serves as an interface to the divider. Thus the goal is to choose an architecture that has minimal impact on the modulation signal integrity — it doesn't restrict its dynamic range and introduces minimal quantization noise.

The simplest structure to choose is a single-stage $\Delta\Sigma$ modulator with a 1-bit quantizer as shown in Figure 3.20 [Cand74]. The full scale input range for such a modulator is

Figure 3.20: Single-stage $\Delta\Sigma$ modulator.

restricted to ± 1 to prevent quantizer overload, but what may not be obvious is its performance for DC values between these limits. For low-order modulators, DC inputs that are integer divisions of the quantizer levels cause limit cycles (i.e. patterns of ones and minus ones) to occur [Gray89]. These limit cycles severely degrade the noise-shaping performance of the modulator leading to a reduction in signal-to-noise ratio, as indicated in Figure 3.21, by the noise peaks within the input range. Operating in these regions is difficult to avoid because the modulating signal could contain a constant stream of ones or zeros (i.e. a DC input that coincides with the susceptible operating region). Higher-order single-bit modulators are more immune to these effects due to the higher degree of randomization of the quantization noise [Cand85],[Ferg90]. However, all the single-stage architectures will exhibit poor noise shaping for signals near their input limits. The reason for this is that the m -bit quantizer in single-stage modulators only uses a few of the available levels to represent an input signal, and near the limits the outer most level is active more often than the others. For example, a $\Delta\Sigma$ modulator with a 1-bit quantizer would represent an input of 0.99 with a constant stream of ones and the occasional minus one. This doesn't provide enough instantaneous activity in the output bitstream to ensure adequate noise shaping, so the in-band noise level increases for signals near the input limits as shown in Figure 3.21. One method of avoiding the poor noise shaping near the input limits is to provide extra quantizer levels beyond what is required. The cascaded architecture in Figure 3.22 achieves this by digitally summing the outputs of two or more $\Delta\Sigma$ modulators into one multi-bit word [Lee87],[Wald90]. The second $\Delta\Sigma$ modulator

Figure 3.21: Noise power of a first-order $\Delta\Sigma$ modulator (OSR=8).

Figure 3.22: Multi-stage (MASH) $\Delta\Sigma$ modulator.

digitally encodes the quantization noise of the first, and subtracts it from the output, leaving only the original signal and the quantization noise of the second $\Delta\Sigma$ modulator. The resulting multi-bit output has a range that exceeds the input range to ensure adequate noise shaping for all input signal levels. The consequence of using a cascaded architecture with k quantizer levels is the need for $\log_2(k + 1)$ $\Delta\Sigma$ FD modulus control bits, where initially one bit was sufficient. However, even with the additional complexity, it is advantageous to use a cascaded architecture since it provides much better performance over the complete input range. A second benefit of cascaded modulator architectures is that they are more stable compared to single-stage modulators of the same order.

3.4.3 Modulation Bandwidth Limitations

Achievable modulation data rates using compensation have so far been restricted by the limited dynamic range of the $\Delta\Sigma$ FD. Other blocks in the synthesizer (e.g. DSP, charge pump etc.) can be altered to handle larger signal swings up to an acceptable limit. The restrictions to date have been circuit oriented but what has yet to be explored is the effect of classic synthesizer loop parameters (e.g. sampling frequency, loop BW, damping factor etc.) on the modulation BW. It was stated in Chapter 2, that the loop parameters are determined by the noise requirements for a given application. Once the noise requirements are met, the design of the Gaussian and equalizer filter can be done independent of the defined loop parameters. However, it is of interest to know what effect these same parameters have on the modulation data rate. The analysis is implementation dependent so some assumptions need to be made. Specifically, the synthesizer order and loop filter need to be chosen to provide adequate attenuation of the quantization noise emanating from the $\Delta\Sigma$ FD. Some arguments were presented earlier that set a lower bound on the order of both the $\Delta\Sigma$ FD and synthesizer. Finally, the sampling frequency is assumed to be much greater than any signal bandwidth (i.e. for a wideband modulator, the modulation data rate defines the maximum BW) so the pseudo-continuous models can be used in the analysis.

3.4.3.1 Effect of Sampling Frequency

It was assumed, in the previous section, that the sampling frequency is much larger than the modulation bandwidth. Based on that assumption, the behaviour of the mixed signal synthesizer approximates that of its continuous-time counterpart for frequencies $f \ll f_r$. However, when wideband modulation is applied with simultaneous equalization, the modulation rate is limited by the input dynamic range of the $\Delta\Sigma$ FD, which in this case is equal to f_r . The immediate solution to this dilemma would be to increase the sampling frequency, thereby increasing the $\Delta\Sigma$ FD input dynamic range by an equal amount. An increased sampling frequency would also relax the DSP requirements due to the increased oversampling ratio (OSR). While this may seem feasible, it turns out that there is a limit to the maximum sampling frequency that ensures correct operation of the $\Delta\Sigma$ FD under steady-state conditions.

Understanding the effect of sampling frequency on the $\Delta\Sigma$ FD operation requires a closer look at the internal signals within the discriminator control loop. A block diagram of the $\Delta\Sigma$ FD is redrawn in Figure 3.23 with the external modulation input included. If the

Figure 3.23: Block diagram of second-order $\Delta\Sigma$ frequency discriminator.

modulation is disabled (i.e. modulus control n is constant), the discriminator loop will force the phase of the divider output to track the reference phase. Due to the coarse single-bit quantization of the phase error that is subsequently fed back through the $(2 - z^{-1})$

block, the divider phase will *bracket* the reference phase with a nominal phase error larger than zero. This differs from a conventional PLL with an integrator loop filter whose steady-state phase error would ideally be forced to zero. In the $\Delta\Sigma$ FD case, it is assumed the peak phase error does not exceed the phase detector dynamic range of $\pm 2\pi$ radians to prevent cycle slipping or else the $\Delta\Sigma$ FD would lose lock. Applying modulation by dithering the divider modulus only aggravates the phase error, since n is summed with the original fed back component b . Both the modulation control n and the fed back component b are $\Delta\Sigma$ noise shaped so the peak value of their sum is the sum of the peak value of each respective signal [John93]. For example, if the modulation control $n \in (100, 101, 102, 103)$ has an average value $\bar{n} = 101.5$ and the fed back component $b \in (0, 1, 2, 3)$ has an average value $\bar{b} = 2$, their sum will have

$$\begin{aligned}\overline{n+b} &= 103.5 \\ n+b &\in (100, 101, 102, \dots 106)\end{aligned}$$

Applying external modulation effectively extends the active modulus range of the divider, leading to a potentially higher peak phase error. Due to the inherent non-linear properties of the $\Delta\Sigma$ FD, it is extremely difficult to calculate the phase error analytically. However, the peak error can be found through simulation of the $\Delta\Sigma$ FD model developed earlier. Since the mean value of n is determined by the channel frequency (modulation dithers n about its mean), the effect of the base modulus on the peak discriminator phase error is of interest. The association between n and the phase error can be found by noting that the error per sample period T_r is

$$\phi_e(k) = 2\pi \left(\frac{T_{div}(k) - T_r}{T_r} \right) \quad [\text{rad}] \quad (3.29)$$

where the divider output period is defined as

$$T_{div}(k) = \frac{n(k)}{f_{vo}} = \frac{n(k)}{(n_o + 1.5)f_r} \quad (3.30)$$

Substituting Equation (3.30) into (3.29) gives

$$\phi_e(k) = 2\pi \left(\frac{n(k)}{(n_o + 1.5)} - 1 \right) \quad [\text{rad}] \quad (3.31)$$

The differential phase error as $n \rightarrow n + 1$ (i.e. $\Delta n = 1$) can be expressed as

$$\Delta\phi_e = 2\pi \left[\left(\frac{n+1}{(n_o + 1.5)} - 1 \right) - \left(\frac{n}{(n_o + 1.5)} - 1 \right) \right] \quad [\text{rad}] \quad (3.32)$$

which reduces to

$$\Delta\phi_e = \frac{2\pi}{(n_o + 1.5)} \quad [\text{rad}] \quad (3.33)$$

From Equation (3.33) it is clear that the phase error will decrease as the base modulus n_o increases as seen in Figure 3.24. The discrepancy between the theoretical and simulated curves exists because the theoretical analysis doesn't consider any non-linear effects. The

Figure 3.24: Effect of $\Delta\Sigma$ FD divider modulus n on PFD peak phase error.

dependence of peak phase error on the divider modulus is intuitively obvious, since a change in $n \rightarrow n + 1$ has much less effect for large values of n . Thus far, the phase error limits have been defined in radians with respect to the reference period T_r . If the reference frequency is increased, the absolute phase error in the time domain will not vary, (i.e. PFD pulse widths remain constant) so the phase error consumes a larger proportion of the phase detectors dynamic range. This effectively sets an upper limit on the reference frequency of the $\Delta\Sigma$ FD, thereby defining its maximum input dynamic range.

3.4.3.2 Synthesizer Loop Bandwidth

Equalization of the synthesizer closed-loop response will only be possible if the chosen loop parameters do not cause the gain of the equalizer to exceed the dynamic range of the synthesizer. In this architecture, the noise and modulation bandwidths can be decoupled and independently set. However, the loop BW cannot be arbitrarily set too low even if this yields the best noise performance. In doing so, the transient response will be compromised (i.e. slower switching speed) and a higher equalizer gain will be necessary to maintain the same modulation BW. Even though the loop BW and modulation BW can be independently set, in practice they are loosely coupled due to the limited dynamic range available.

3.4.4 Effect of Mismatch

Efforts to equalize the effects of the synthesizer closed-loop transfer function have thus far been based on the assumption that the synthesizer closed-loop response is well defined. The main advantage of this modulator architecture over the one described in [Perr97] is that all the loop parameters except for open-loop gain are predictable since most of the synthesizer is digital. This offers two immediate benefits:

- loop parameters are insensitive to process and temperature variation
- it is easy to digitally equalize a transfer function that is mostly discrete time

The remaining issue to resolve is the effect of the open-loop gain on the performance of

the modulator. There are two cases to consider — the effect on the synthesizer stability and the effect on the modulation path.

3.4.4.1 Loop Stability

To determine how a change in the open-loop gain K will alter the loop stability, a review of the basic loop dynamics is necessary. The equations derived earlier for the synthesizer loop parameters are restated below.

$$\begin{aligned} BW &= K'(a_2 + a_3) \\ \omega_n &= \sqrt{-K' f_r \ln \left[\left(\frac{a_2}{a_2 + a_3} \right)^{(a_2 + a_3)} \right]} \\ \zeta &= \frac{\sqrt{K'(a_2 + a_3)}}{2 \sqrt{-f_r \ln \left(\frac{a_2}{a_2 + a_3} \right)}} \end{aligned}$$

From these equations, it is clear that the loop bandwidth is proportional to the open-loop gain K and the natural frequency ω_n and damping factor ζ both vary as $\sqrt{K}$. Table 3.1 shows the change in loop parameters with a $\pm 20\%$ change in open-loop gain. Since the

Table 3.1: Effect of open-loop gain error on loop parameters.

gain error	BW	ω_n	ζ
-20%	24 KHz	9.9 KHz	1.2
0%	30 KHz	11 KHz	1.4
+20%	36 KHz	12.1 KHz	1.5

compensating zero in the loop filter does not depend on the open-loop gain while loop BW does, the stability of the loop will vary with open-loop gain.

3.4.4.2 Open-Loop Gain Error

The path that the modulation data is subjected to is shown graphically in Figure 3.25, where the synthesizer has been replaced by the equivalent modulation closed-loop transfer

Figure 3.25: Modulation data path.

function. Ideally, the equalizer exactly compensates for the closed-loop attenuation within the equalization BW (i.e. accept attenuation at frequencies beyond this) but a gain error will cause an over or under compensation of the modulation transfer function. This is evident in Figure 3.26 where the normalized equalization transfer function is correct

Figure 3.26: Misshaped modulation transfer function due to open-loop gain error.

within the synthesizer loop BW. At frequencies near the loop BW, there is a noticeable peak (and trough) which is due to the second-order closed-loop transfer function, and beyond this the transfer function exhibits a relatively flat gain error. The influence of the equalization mismatch in the time domain is discussed in Chapter 5, where the demodulated GSM signals yield eye diagrams that deviate from the ideal GSM modulation.

3.5 Gaussian Pulse Shaping

Pulse shaping is used in a transmitter to reduce the spectral BW of the RF signal. Without it, the rectangular data pulses would cause the theoretical power spectrum to spread infinitely wide, although in practice the spectral power attenuates with a $\text{sinc}(f)$ response. In this modulator, the modulation scheme is restricted to GMSK which uses Gaussian filtering to constrain the spectrum. It is worthwhile exploring the characteristics of Gaussian filters with various symbol data rates and normalized filter bandwidths BT to identify the effect on the modulation BW. The argument for this is that the filtered modulation data has to pass through the digital $\Delta\Sigma$ modulator and the $\Delta\Sigma$ FD, both of which utilize noise shaping. If the modulation BW is too wide, the shaped quantization noise power would dominate at high frequencies leading to a poor overall SNR. Equalization provides some redeeming qualities, since compensation for the synthesizer closed-loop transfer function amplifies the high frequencies of the modulating signal, which effectively raises the SNR (i.e. the absolute noise floor doesn't matter as long as the signal is larger).

Figure 3.27(a) compares the effect of varying the filter bandwidth BT with a fixed data rate while Figure 3.27(b) shows the effect of varying the symbol rate with a constant filter bandwidth. If the usable energy of the modulation signal is restricted to frequencies from DC to where the sidelobes are suppressed to a certain level, Figure 3.27 shows that

Figure 3.27: GMSK baseband modulation bandwidth with varying (a) filter bandwidth BT and (b) symbol rate.

the symbol rate has more impact on BW than the Gaussian filter normalized bandwidth. However, it should be noted that even a relatively small symbol rate with a large BT can result in a large modulation bandwidth.

Implementation of the Gaussian filter using a ROM based FIR filter is desirable in the sense that the filter output is the result of a simple table look-up and doesn't require any arithmetic operations. The content of the ROM is simply the pre-computed trajectories for all the possible input data sequences. The key issue to remember is that narrow filter bandwidths result in long impulse responses (i.e. partial response system where filter impulse response extends over several data symbol periods). Since the tap weights of a FIR filter are identical to its impulse response, a small BT leads to a more complex digital filter.

3.6 Digital $\Delta\Sigma$ Modulator

The $\Delta\Sigma$ modulator in the modulation path provides an interface between the filtered modulation data and the $\Delta\Sigma$ FD modulus control input n . If the $\Delta\Sigma$ FD modulus control had

a large enough dynamic range, the digital $\Delta\Sigma$ modulator wouldn't be necessary. However, there are practical limits to the range of n that can be implemented, which is generally much less than that required by the Gaussian filtered and equalized modulation data. The goal of the $\Delta\Sigma$ modulator is to coarsely interpolate the modulation data without introducing too much in-band quantization noise. That is to say, convert a k -bit word into an m -bit word where $m < k$, while simultaneously shaping most the quantization noise power out of band, where it will be filtered by the synthesizer closed-loop transfer function.

The validity of using a cascaded modulator architecture to make use of the entire input dynamic range, while maintaining an adequate SNR, was discussed earlier. What remains to be resolved is choosing an appropriate modulator order that is compatible with the rest of the synthesizer. The key issue here is that the $\Delta\Sigma$ modulator quantization noise will be filtered by the *existing* synthesizer closed-loop transfer function seen by the modulation data. A second-order synthesizer was previously chosen to ensure adequate filtering of the $\Delta\Sigma$ FD quantization noise so the order of the digital $\Delta\Sigma$ modulator must be equal to or less than two. Maximum modulation quantization noise shaping is achieved by using a $\Delta\Sigma$ modulator with the highest allowable order (with respect to the synthesizer order) so a second-order architecture is chosen. The quantization noise powers of the $\Delta\Sigma$ FD and digital $\Delta\Sigma$ modulator simply add, since they are uncorrelated. Since both the $\Delta\Sigma$ FD and digital $\Delta\Sigma$ modulator have the same order, the noise floor will raise by 3dB, which is tolerable as long as the overall SNR remains satisfactory.

A second-order cascaded (MASH) architecture shown in Figure 3.28, will satisfy the dynamic range requirements and its quantization noise, along with the noise from the $\Delta\Sigma$ FD, will be adequately filtered by the synthesizer closed-loop response. The MASH

Figure 3.28: Digital second-order MASH $\Delta\Sigma$ modulator.

modulator consists of two first-order $\Delta\Sigma$ modulators whose outputs are combined to give the final extended range multi-bit value. The error from the first modulator is encoded, differentiated and digitally subtracted from the output, leaving only the original signal and the error from the second modulator. From Figure 3.28, Z-domain expressions for the signal and noise transfer functions are

$$\text{signal T.F.} = z^{-2}$$

$$\text{noise T.F.} = (1 - z^{-1})^2$$

where the signal has a flat response, albeit delayed, and the quantization noise of the second stage (the first stage has its noise cancelled) is second-order shaped. The $\Delta\Sigma$ modulator dynamic range is determined by the number of levels in the quantizers. In this application, they can be simple 1-bit quantizers (i.e. comparators) since the dynamic range of the $\Delta\Sigma$ FD is limited to f_r which corresponds to an *average* modulus range $(n, n + 1)$. Note that the *instantaneous* modulus range is $(n - 2, n - 1, \dots, n + 2)$, due to the MASH architecture so the $\Delta\Sigma$ FD modulus control must have at least $\lceil \log_2(5) \rceil$ bits of dynamic range.

A new wideband modulator architecture has been described that uses digital equalization to compensate for the limited closed-loop BW of a $\Delta\Sigma$ FD based synthesizer.

Potential mismatch between the synthesizer closed-loop response and the equalizer response is minimized since most of the synthesizer is digital. This eliminates any analog process and temperature variations that would occur in conventional PLL architectures. In Chapters 4 and 5, implementation details of a GMSK modulator that is suitable for wireless digital radio applications are addressed.

Chapter 4

A 2.5 GHz BiCMOS $\Delta\Sigma$ Frequency

Discriminator

The $\Delta\Sigma$ FD is a central component in this modulator architecture and it also forms a control loop within the main synthesizer loop. Therefore, it is imperative to independently explore its performance from an architectural perspective and then use these results to design an integrated realization suitable for use in the GMSK modulator. The key to understanding the discriminator operation is to form an analogy with conventional $\Delta\Sigma$ modulators, whose operation and performance are well known [Cand85]. Once this is accomplished, it is possible to determine the expected performance of various architectures under ideal conditions and while being influenced by non-ideal effects.

The analyses in Section 4.1 concentrate on the second-order single-loop $\Delta\Sigma$ FD architecture illustrated in Figure 4.1. This discriminator architecture is used in the final modulator although the results apply to any n^{th} -order structure. The goal in this section is to quantify any fundamental limitations that pertain to the proposed single-loop discriminator architecture. Section 4.2 describes the design of the BiCMOS frequency discriminator chip, whose realization is based on the results obtained from the upcoming architectural analysis. Various implementation strategies are used to achieve good high frequency performance while curtailing power consumption wherever possible.

4.1 $\Delta\Sigma$ Frequency Discriminator Architecture

Figure 4.1: Single-loop, second-order $\Delta\Sigma$ frequency discriminator.

For convenience, the block diagram of the single-loop second-order $\Delta\Sigma$ FD is redrawn in Figure 4.1. The first step in analyzing this circuit is to make use of the model developed in Section 3.2 to develop an analogy between $\Delta\Sigma$ frequency discriminators and

Figure 4.2: Second-order single-loop $\Delta\Sigma$ frequency discriminator model.

conventional $\Delta\Sigma$ modulators. Substituting each block in Figure 4.1 with its equivalent model yields the linear discriminator model in Figure 4.2. Note that the multi-modulus divider has been replaced by its equivalent model developed in Section 3.2.1. This step is necessary to readily identify the resemblance of this architecture to that of a conventional second-order $\Delta\Sigma$ modulator. The key difference is the definition of the input variables which in the case of the $\Delta\Sigma$ FD are the RF input frequency f_v and modulation control n . Assuming a 1-bit quantizer with an effective gain K_q , an expression for the average output bitstream that ignores all non-ideal noise effects (except quantization noise) is

$$f_{out} = \frac{1}{f_r} f_v - n + (1 - z^{-1})^2 S_Q \quad (4.1)$$

where the 1-bit quantizer gain absorbs all the constants in the loop and the *average* divider modulus $N = f_{vo}/f_r$. If the modulation input n is fixed, its effect on the output bitstream will only be a constant DC offset. The remaining terms in Equation (4.1) indicate that the signal transfer function is uniform but scaled by one over the reference frequency f_r , and the quantization noise S_Q is second-order shaped. The noise transfer

Figure 4.3: Second-order frequency discriminator: (a) noise transfer function and (b) pole-zero plot.

function $(1 - z^{-1})^2$ arises from the implicit conversion of frequency into phase in the divider and the analog charge pump in the forward path. This double integration results in a lowpass noise transfer function as shown in Figure 4.3(a) or alternatively a pole-zero plot with two poles at DC and two zeros at $z = 1$. From Figure 4.3 it becomes clear that the noise S_Q introduced by quantizing the signal is shaped through the feedback action of the loop. The choice of the error filter transfer function, which in this case is two integrators, is somewhat arbitrary for conventional $\Delta\Sigma$ modulators. In the case of the second-order $\Delta\Sigma$ FD, the first filter is an implicit integrator (due to the conversion of frequency into phase in the divider) so are no circuit imperfections to worry about. The second filter is also an integrator but realized as an analog charge pump. Due to the analog realization, circuit imperfections will alter the response to less than ideal. Note that any filter transfer function could be used for the second filter including one with a non-monotonic response (e.g. a resonator). However, unlike in conventional $\Delta\Sigma$ modulators, the second filter must be able to tolerate discrete pulses from the PFD and a charge pump is best suited for that purpose.

Noise shaping and oversampling offers a clear advantage over simply quantizing the signal at the Nyquist rate within a limited bandwidth as indicated in Figure 4.4. This is

Figure 4.4: Comparison between Nyquist, oversampled and $\Delta\Sigma$ noise shaped quantization.

characteristic of $\Delta\Sigma$ modulators which use oversampling and noise shaping to shift most of the inherent quantization noise power out of the frequency band of interest. Note that the total noise power remains constant for all three cases, but by noise shaping the

oversampled signal and filtering the undesired out-of-band noise, the in-band signal-to-noise ratio is improved. This post filtering operation can be realized in a number of ways depending on the application, but when analyzing a modulator architecture it is simply assumed to impose some form of band limiting of the total signal including noise.

4.1.1 Non-linear Effects

The analysis thus far has been based on the linear model developed in Chapter 3. However, it is clear that the $\Delta\Sigma$ FD, or any $\Delta\Sigma$ modulator, is inherently non-linear due to the quantizer, so modelling it as a linear system cannot reveal its true characteristics although it does provide useful results. Therefore time-domain simulation using a non-linear model will be used in the following sections to extract the true $\Delta\Sigma$ FD characteristics. Bear in mind that non-linear characteristics do not imply non-ideal effects due to circuit limitations, since non-linear effects exist even in the ideal $\Delta\Sigma$ FD or $\Delta\Sigma$ modulator. The non-linear time-domain model illustrated in Figure 4.5 uses abstract blocks that replicate

Figure 4.5: Non-linear SIMULINK[®] model used for time-domain simulation.

the functionality of the discriminator without resorting to an actual circuit implementation. The reason for this is twofold — the simulation time is much less when using abstract models and the actual circuit implementation is not known at this time.

The most prominent difference between results obtained from linear and non-linear models is a direct result of the assumption that a quantizer can be modeled as an ideal gain with an additive uncorrelated white noise source [Benn48]. While there are cases where both models may produce similar results, it is generally not a valid assumption to make, since the quantization noise is signal dependent and not truly uncorrelated. A classic example of correlated quantization noise occurs when a $\Delta\Sigma$ modulator operates with DC inputs (constant input frequency for $\Delta\Sigma$ FD's). The quantized signal alternates between two levels, keeping the mean value equal to the DC input level. This oscillation may be repetitive which results in a pattern that repeats after a number of samples leading to distinct idle tones in the output spectrum. If the sequence length is long enough, tones will exist within the signal BW and degrade the overall SNR. Inputs that are integer divisions

Figure 4.6: Noise power of a first-order $\Delta\Sigma$ modulator (OSR=8).

of the quantizer level spacing Δ (e.g. 1/4, 1/3, 1/2 etc. for $\Delta = 1$) cause the harmonics to alias exactly to DC. This has no effect on the total noise power since these tones are subsequently filtered. Figure 4.6 compares the total noise power for various DC inputs and clearly shows the nulls where the limit cycle tones are aliased to DC. On either side of these input levels, the total noise power is much larger due to the power of the in-band tones. The existence of in-band tones for certain DC inputs seems to restrict the use of a $\Delta\Sigma$ modulator since its SNR will be severely degraded, but one can reduce this problem in a number of ways:

- ensure the input signal isn't DC (i.e. there is a large enough AC component)
- randomize the modulator states by adding some dither to the input or internal states

Randomizing the input signal (or internal states) doesn't eliminate the in-band tone problem but it simply prevents the modulator from operating *continuously* in this region. In the case of the GMSK modulator presented here, the GMSK modulation data is injected into the $\Delta\Sigma$ FD input n , which ensures that the discriminator input is kept busy enough to prevent idle tones even though the frequency input (i.e. channel) is kept constant.

There is an additional phenomenon that occurs when a $\Delta\Sigma$ modulator operates with a slowly changing DC input (i.e. a frequency ramp for $\Delta\Sigma$ FD's). In this case, the output of the integrator (see Figure 4.5 for the equivalent frequency discriminator block diagram) can shift some amount between two adjacent quantizer input threshold levels without any effect on the quantized output. Such a change of level at the integrator output corresponds to an impulse at its input. Consequently, small fast changes in the $\Delta\Sigma$ modulator input may be ignored under certain conditions, leading to transient deadzones in the average transfer function of the modulator as seen in Figure 4.7. Within the deadzone, the $\Delta\Sigma$ modulator settles into a periodic pattern and ignores the input signal. It can be shown that the deadzone regions correspond to the noise peaks of Figure 4.6. For most applications, the idle tones caused by limit cycles are more noticeable than the effect of the deadzone, but

Figure 4.7: Deadzone effect in a second-order $\Delta\Sigma$ modulator with integrator leakage (gain=64) and OSR=64.

this may be reversed if the integrator has low DC gain (i.e. a leaky integrator).

From an architectural perspective, a multi-level or single-level quantizer may be used for the second-order $\Delta\Sigma$ FD in Figure 4.1. Since quantizing a signal is inherently a non-linear operation, one may be tempted to use as many levels as possible. This will result in a total quantization noise power

$$S_{Q(RMS)}^2 = \frac{\Delta^2}{12} \quad (4.2)$$

where Δ is defined as the quantizer range divided by the number of levels. However, it is more difficult to design a multi-level quantizer because any misplaced threshold levels may be regarded as a non-linearity in gain, where gain is the output level spacing divided by the input threshold spacing. This may not be a significant problem in this $\Delta\Sigma$ FD architecture because the quantizer is preceded by two high gain integrators so a slight gain error is tolerable. In conventional $\Delta\Sigma$ modulators, a multi-level quantizer is accompanied

by a D/A converter with equal range in the feedback path. Errors in the feedback path can not be corrected, so the linearity of the D/A converter must be quite high. This is the reason that multi-level quantization is difficult to implement in conventional $\Delta\Sigma$ modulators. For the single-loop $\Delta\Sigma$ FD architecture, the advantage is that the feedback path is entirely digital so linearity is perfect as opposed to the multi-loop $\Delta\Sigma$ FD architecture described in [Bear94]. In theory, the single-loop architecture could use any number of quantization levels but practically it is limited by the range of low-delay moduli b (i.e. those controlled by the feedback path) available in the multi-modulus divider. These moduli differ from the conventional modulus input n used to set the channel offset since they are allowed to be modified during the current divide cycle. The reason for the low-delay moduli is explained in the implementation section that follows. This limitation is further compounded by the $(2 - z^{-1})$ feedback logic which increases the dynamic range of the feedback signal. For example, if a 1-bit quantizer (i.e. two levels) is used, the $(2 - z^{-1})$ feedback logic would produce values ranging from zero to three. This implies that the multi-modulus divider must be able to dynamically divide by one of four possible values. The required low-delay modulus range of the divider with increasing quantizer levels (assuming a constant range) is shown in Table 4.1. Beyond two quantizer levels (i.e.

Table 4.1: Divider modulus range for various quantizer resolutions.

quantizer levels	modulus range
2	4
4	16
8	32
16	64

1-bit), the architecture of the multi-modulus divider rapidly becomes more complex which is why a two level quantizer is used in this $\Delta\Sigma$ FD implementation. An added benefit of using a simple two level quantizer is that it is inherently linear since its gain is undefined. Thus the single-loop $\Delta\Sigma$ FD architecture has a clear advantage, since there are no quantizer related linearity problems in the forward path or any non-linearities in the digital feedback path.

4.1.2 Enhancing the Input Sensitivity

The sensitivity of the $\Delta\Sigma$ FD, without considering circuit imperfections, is directly related to its dynamic range and determines the minimum detectable input frequency. If the input frequency deviation consumes a large part of the $\Delta\Sigma$ FD dynamic range, it will be able to faithfully produce a demodulated version of the RF input since the SNR will be high. Here the assumption is that the only noise is due to quantization and it remains constant for a given signal BW. For the single-loop $\Delta\Sigma$ FD described earlier, the input dynamic range is defined as $(n+1)f_r < f_{in} < (n+2)f_r$ due to the $(2-z^{-1})$ in the feedback path. This yields a dynamic range equal to the sampling frequency f_r . The average value of the oversampled digital output is a measure of the instantaneous frequency deviation of the carrier from the nominal value $(n+1.5)f_r$. Thus for a 1-bit quantizer producing a zero or a one, the average output becomes

$$\overline{f_{out}(t)} = 0.5 + \frac{1}{f_r} [f_{in}(t) - (n(t) + 1.5)f_r] \quad (4.3)$$

where $f_{in}(t) = f_c + \Delta f(t)$ is the modulated carrier and n is the base modulus of the divider. If the input signal has a maximum deviation of $\pm\Delta f_{max}$, the average output bitstream has a corresponding range

$$0.5 - \frac{\Delta f_{max}}{f_r} < \overline{f_{out}} < 0.5 + \frac{\Delta f_{max}}{f_r} \quad (4.4)$$

This results in poor use of the available dynamic range of the $\Delta\Sigma$ FD when Δf_{max} is small (i.e. $\Delta f_{max} \ll f_r/2$). Reducing f_r to permit the input signal to occupy a greater proportion of the dynamic range is undesirable since f_r determines the oversampling ratio of the $\Delta\Sigma$ FD and ultimately the transmit modulation bandwidth. On the other hand, the input signal is angle modulated with a fixed maximum frequency deviation which cannot be altered. However, the $\Delta\Sigma$ FD reacts to the frequency difference between the input frequency and the centre frequency $(n+1.5)f_r$. If one were to vary the base modulus n over time, the $\Delta\Sigma$ FD would simply produce an output that corresponded to the instantaneous frequency deviation of the input from the now time-varying centre

frequency $(n(t) + 1.5)f_r$. If the value of n were to track the input frequency deviation $\Delta f(t)$, it would alter the effective deviation that the $\Delta\Sigma$ FD sees. For example, if n varies from its nominal value by $k\Delta f(t)$ where k is a constant, the $\Delta\Sigma$ FD output becomes

$$\overline{f_{out}}(t) = 0.5 + \frac{1}{f_r} [f_{in}(t) - (n + k\Delta f(t) + 1.5)f_r] \quad (4.5)$$

The corresponding output range is

$$0.5 - \frac{\Delta f_{max}(1 - kf_r)}{f_r} < \overline{f_{out}} < 0.5 + \frac{\Delta f_{max}(1 - kf_r)}{f_r} \quad (4.6)$$

and if $k = -\alpha/f_r$, the range becomes

$$0.5 - \frac{(1 + \alpha)\Delta f_{max}}{f_r} < \overline{f_{out}} < 0.5 + \frac{(1 + \alpha)\Delta f_{max}}{f_r} \quad (4.7)$$

Note that the effective frequency deviation seen by the $\Delta\Sigma$ FD has been increased by a factor of $(1 + \alpha)$ by allowing n to be a function of the input frequency deviation $\Delta f(t)$. The divider modulus is now composed of the original constant term n and a time varying component that tracks the frequency deviation of the RF input signal as shown in Equation (4.8).

$$n(t) = n - \frac{\alpha\Delta f(t)}{f_r} \quad (4.8)$$

Although this technique seems like a simple way to improve the input sensitivity, there are two issues to resolve before it can be used. First, the $\Delta\Sigma$ FD is a sampled system while $n(t)$ is continuously time varying and second, the divider modulus n must be an integer while $n(t)$ (due to the α term) can attain any arbitrary value. The first problem can be solved by oversampling $\Delta f(t)$ to get $\Delta f(k)$, which is simply the output bitstream of the $\Delta\Sigma$ FD sampled at f_r . Quantizing the continuous-value (high resolution) modulus into integers can be accomplished by remodulation using a $\Delta\Sigma$ modulator. Thus the divider modulus becomes

$$n(k) = n - \frac{\alpha \Delta f(k)}{f_r} \quad (4.9)$$

The time varying component of $n(k)$ in Equation (4.9) is the input frequency deviation scaled by $-\alpha/f_r$. However, a measure of the frequency deviation can be extracted from the $\Delta\Sigma$ FD output, which is the input deviation scaled by $(1 + \alpha)$ and normalized with respect to f_r . Thus the required feedback gain, assuming the signal gain of the $\Delta\Sigma$ modulator in the feedback path is one, becomes

$$\text{gain} = \frac{-\alpha/f_r}{(1 + \alpha)/f_r} = \frac{-\alpha}{1 + \alpha} \quad (4.10)$$

The modified $\Delta\Sigma$ FD architecture shown in Figure 4.8 is the same as the original one in Figure 4.1 (shaded region), except for an extra feedback path containing a digital $\Delta\Sigma$

Figure 4.8: Modified single-loop $\Delta\Sigma$ frequency discriminator block diagram.

modulator.

In the context of the wideband GMSK transmitter, which uses direct modulation of a synthesizer, detecting small deviation angle modulated signals is not an issue so the

original architecture in Figure 4.1 is used. The reason input sensitivity is not a problem is due to the fact that equalization is used to compensate for signal attenuation at high modulation frequencies (with respect to the loop PLL BW). The equalized modulation signal has a much larger dynamic range than the original GMSK signal and it is injected into the $\Delta\Sigma$ FD's modulation input n . As described earlier, the $\Delta\Sigma$ FD detects the difference between the RF input frequency (i.e. VCO output) and its centre frequency $(n(t) + 1.5)f_r$ and produces a bitstream whose average density is

$$\overline{f_{out}(t)} = 0.5 + \frac{1}{f_r} [f_{in}(t) - (n(t) + 1.5)f_r] \quad (4.11)$$

The VCO produces a GMSK modulated carrier with the desired frequency deviation but the base modulus n is now being controlled by the equalized modulation signal. The *difference* between these two signals exercises much more of the available dynamic range in the discriminator, even though the transmitter output signal contains the original frequency deviation.

Although this technique is not implemented in the GMSK modulator, it was used in a GSM receiver to increase the effective input sensitivity of a $\Delta\Sigma$ FD to detect narrowband angle-modulated signals [Bax98b]. Extraction of the modulation data was accomplished by filtering the $\Delta\Sigma$ FD output bitstream using a digital matched filter with a response identical to that in Figure 4.9. This removes most of the quantization noise and minimizes the mean square error (MSE) or alternatively, maximizes the signal-to-noise ratio. What remains is the original GMSK modulation expressed in the form of eye diagrams as in Figure 4.10. For comparison, Figure 4.10(a) illustrates the received eye diagram of the original GSM signal without input sensitivity enhancement, while Figure 4.10(b) shows the enhanced eye diagram obtained when the extra feedback path in Figure 4.8 is enabled. There is a slight degradation of the enhanced eye opening due to the additional quantization noise added by the digital $\Delta\Sigma$ modulator in the feedback path. However, the eye opening is clearly enlarged at the sampling point (i.e. maximum eye opening) which increases the noise margin during the decision making process, leading to a lower bit error rate.

Figure 4.9: Digital matched filter response for GSM modulation.

Figure 4.10: Filtered GSM eye diagram from (a) original and (b) modified $\Delta\Sigma$ FD output.

4.1.3 Loop Stability

Earlier stability analysis of the main synthesizer loop was based on the assumption that the $\Delta\Sigma$ FD itself was inherently stable and could be represented by an equivalent stable linear system. In this section, the focus will be on whether the proposed $\Delta\Sigma$ FD architecture can be shown to be stable in the context of $\Delta\Sigma$ modulators (i.e. stability implies that they exhibit chaotic behaviour).

Stability analysis of $\Delta\Sigma$ modulators differs from conventional linear system analysis due to the existence of the quantizer, which is inherently non-linear. Thus it is not surprising, that attempting to use linear system stability analysis often fails, although it sometimes provides valuable insight for low-order modulators [Arda87]. The approach that will be used here is to use linear stability analysis on the discriminator model developed in Chapter 3 to determine the parameters that ensure stability. This gives some insight into the expected performance of the discriminator but it by no means guarantees stability for all input conditions. Further proof of stability is assured by simulating the

Figure 4.11: Simplified second-order $\Delta\Sigma$ frequency discriminator linear stability model.

non-linear model of the $\Delta\Sigma$ FD (from Figure 4.5) with various input conditions to verify that the architecture has bounded states and is indeed stable.

Figure 4.11 depicts the linear model used in the stability analysis. It is the same as the one developed in Chapter 3 without the lumped noise sources except for the linear model of the quantizer with its associated gain K_q . The dilemma arises from the fact that the 1-bit quantizer has an undefined gain (its output is the sign of its input). This implies that the overall loop gain K is also undefined and further analysis isn't possible. One solution is to model the 1-bit quantizer with a gain that minimizes the error signal's power. There is an optimum value for K_q which results in the error and signal components being uncorrelated and is given by

$$K_q = \frac{cov(x, y)}{var(y)} = \lim_{N \rightarrow \infty} \frac{\sum_{n=0}^N x(n)y(n)}{\sum_{n=0}^N x(n)^2} \quad (4.12)$$

where x and y are the input and output of the quantizer respectively [Arda87]. This formula clearly shows that K_q depends on the quantizer input which in turn depends on the discriminator input. The linear quantizer model in Figure 4.12 is still the same, but the

Figure 4.12: Linear signal-dependent model of a 1-bit quantizer.

gain is now signal dependent. Since K_q forms part of the overall loop gain, it is necessary to determine its value for various input signal levels so the stability analysis can continue. To do so, the non-linear discriminator model was simulated with various DC input levels (i.e. constant carrier frequency). In each case, the AC components of the actual quantizer

input and output were determined and a value for K_q was computed according to Equation (4.12). These results are shown in Figure 4.13 where the DC level is normalized

Figure 4.13: Signal dependent gain of a 1-bit quantizer (modulus $n=142$).

to the offset from the centre frequency. Only one half off the full range is shown since the same effect occurs for $f_{in} < f_{mid}$. From this plot one can see that the quantizer gain drops as the DC level increases and has a peak value at the centre frequency where $DC=0$. Note that the inclusion of the divider in the loop forces the effective quantizer gain K_q to attain a value other than one for DC inputs near zero (K_q approaches the divider modulus n as DC input is reduced to zero) as compared to conventional $\Delta\Sigma$ modulators.

Having defined the value of K_q for all input signal levels, it becomes possible to determine the loop stability. From the model in Figure 4.11, the open-loop gain $G_{OL}(z)$ is defined as

$$\frac{K_q f_{vo}}{N f_v} \cdot \frac{(2z^{-1} - z^{-2})}{(1 - 2z^{-1} + z^{-2})} \quad (4.13)$$

where the total loop gain is

$$K = \frac{K_q f_{vo}}{N f_v} \quad (4.14)$$

Determining the closed loop stability amounts to solving for $1 + G_{OL}(z) = 0$ for various loop gains and checking to see if the poles are within the unit circle of the Z-plane. Results

Figure 4.14: Root-locus plot of the linear $\Delta\Sigma$ FD model.

of these computations are plotted in Figure 4.14 as a root-locus plot in the Z-plane. From this plot, it is clear that the discriminator is stable (in a linear sense) for loop gains ranging from $K = 0 \rightarrow 1.33$. From Equation (4.14), the corresponding quantizer gain should be constrained so $K_q < 1.33N$ where N is the average divider modulus. This is true for all DC input levels assuming a narrowband input signal (i.e. $f_v \approx f_{vo}$) according to Figure 4.11. While linear stability analysis seems to impose some limit on the loop gain, there really isn't a problem since the 1-bit quantizer responds to the sign, not the amplitude of its input.

Even though the discriminator loop is stable for a wide range of loop gains, its

performance is adversely affected by the signal-dependent quantizer gain. There are two cases to consider — the effect on the signal and the quantization noise. From Figure 4.11, an expression for the signal gain is

$$H_{SIG}(z) = \frac{NK_q}{Nf_v + (2K_qf_{vo} - 2Nf_v)z^{-1} + (Nf_v - K_qf_{vo})z^{-2}} \quad (4.15)$$

A plot of the signal response as the quantizer gain varies with DC input levels is shown in Figure 4.15. For the nominal midband case where $K_q = Nf_v/f_{vo}$, the signal gain has a flat response with a gain of -142.3dB. This matches the expected signal scaling factor of

Figure 4.15: $\Delta\Sigma$ FD signal transfer function for various quantizer gains.

$1/f_r$ in the linear model where $f_r = 13\text{MHz}$. However, as the DC level (i.e. carrier frequency) is altered, the signal transfer function begins to peak followed by an increasing rate of attenuation. This will only effect the input signal if it has a large frequency deviation. If the $\Delta\Sigma$ FD was constrained to always operate at or near its centre frequency (DC input is zero), distortion of the signal can be avoided. This is exactly the case for the

GMSK modulator since controlling the modulation port n forces the discriminator to operate at its midband.

Using the same linear model as before, the quantization noise can be expressed as

$$H_{NOISE}(z) = \frac{Nf_v(1 - z^{-1})^2}{Nf_v + (2K_q f_{vo} - 2Nf_v)z^{-1} + (Nf_v - K_q f_{vo})z^{-2}} \quad (4.16)$$

Operation of the discriminator at its centre frequency implies that the quantizer gain is $K_q = Nf_v/f_{vo}$ and Equation (4.16) reduces to the familiar form

$$H_{NOISE}(z) = (1 - z^{-1})^2 \quad (4.17)$$

representing ideal second-order double integration noise shaping. As the input carrier frequency shifts away from the centre point, the noise shaping begins to peak and usable BW progressively decreases as shown in Figure 4.16. For a given signal BW, the total

Figure 4.16: $\Delta\Sigma$ FD noise transfer function for various quantizer gains.

noise power will start to increase resulting in a reduced SNR.

It has been shown in [Stein94], that second-order $\Delta\Sigma$ modulators with integrator filters are stable for all DC inputs properly bounded by the range of the quantizer. That is to say, the internal states of the modulator (i.e. the two integrator outputs) are bounded for all initial conditions and the quantizer recovers from temporary overload. However, modulator stability for some AC signals with amplitudes constrained by the same limits is not necessarily guaranteed, implying that the second-order modulator is not unconditionally stable for arbitrary AC inputs. Analytical proof of $\Delta\Sigma$ modulator stability for arbitrary AC inputs is usually not feasible (if not impossible), so simulation is used instead using expected input signals. Analytical non-linear stability analysis of the second-order $\Delta\Sigma$ FD is impractical for the same reasons, so the same simulation techniques are used to ensure stability.

The $\Delta\Sigma$ FD differs from conventional $\Delta\Sigma$ modulators in that the input frequency is the variable quantity rather than amplitude. The theoretical input frequency dynamic range for the single-loop structure is $(n+1)f_r \rightarrow (n+2)f_r$, as defined by the reference frequency f_r . Restricting the input signal to DC levels is analogous to injecting an RF carrier without modulation such that $(n+1)f_r < f_{in} < (n+2)f_r$. What remains to be done is to identify the two integrator outputs in the actual circuit to ultimately plot the state space diagram. The second integrator is clearly the output of the charge pump while the first integration is implicitly performed in the phase-frequency detector. During steady-state conditions (i.e. no cycle slipping), the PFD operates as a phase detector and generates a pulse whose width represents the phase error of the current sample. The integration occurs in the implicit conversion of frequency into phase since $\phi = 2\pi \int f(t)dt$. Converting the PFD output pulse width into the equivalent phase error (i.e. integral of frequency error) gives

$$\phi(k) = 2\pi \frac{t_{pulse}(k)}{T_r} \quad [\text{rad}] \quad (4.18)$$

As in conventional $\Delta\Sigma$ modulators, clipping of the integrators must be avoided to prevent saturation limit cycles [Bair94]. In this architecture, the limits of the first integrator are set

by the linear dynamic range of the PFD (typically $\pm 2\pi$ radians) while the second integrator output must remain within the compliance range of the charge pump. State space diagrams for three test cases with DC inputs (i.e. constant RF frequency) are shown in Figure 4.17. The plots in Figure 4.17(a) and 4.17(c) are parabolic in shape due to the

Figure 4.17: $\Delta\Sigma$ FD state space diagram for (a) low frequency, (b) midband frequency and (c) high frequency RF input signals.

two integrators in the loop and indicate that the quantizer does temporarily overload (i.e. 1-bit quantizer input exceeds ± 1) but it recovers within a few samples. The $\Delta\Sigma$ FD is stable with these inputs but the noise shaping will be severely degraded due to quantizer overload. The plot in Figure 4.17(b) shows normal operation for an RF input frequency *near* midband (exact midband operation with DC inputs causes limit cycles under ideal conditions although practically, circuit noise randomizes the periodic output).

4.1.4 Acquisition

Section 4.1.3 described various approaches to ensure that the $\Delta\Sigma$ FD is stable under various steady-state operating conditions. What wasn't mentioned was how the discriminator handles an input signal with various initial conditions. That is to say, can the loop lock for all initial states of the integrators? The important point to remember is that there are two inherent non-linearities present in the loop whose effect will alter initial acquisition. The

first non-linear element to consider is the quantizer which has already been discussed. In the context of loop acquisition, the effect of the quantizer is to distort the true output of the second integrator which ultimately is fed back to control the multi-modulus divider. What can be assured is that the polarity of the quantized output is correct for all quantizers, and for multi-bit quantizers, the magnitude is increasingly more accurate for those with higher resolution. Now the second non-linear element is the first integrator which is implemented as a digital phase-frequency detector (PFD). As long as the phase error between the reference and divider output frequency is within the linear range of the PFD, the correct output pulse will be produced and normal operation continues. However, if there is an initial frequency error (i.e. $f_r \neq f_{div}$), it will cause the phase error to grow since $\phi_{err}(t) = 2\pi \int f_{err}(t) dt$. If this condition is allowed to continue, the phase error will exceed the $\pm 2\pi$ linear range of the PFD and cycle slipping will occur. This effect is identical to what happens in conventional phase-locked loops (PLL's) with digital PFD's during their acquisition phase. The key point in acquiring lock is that while the PFD may be cycle slipping in phase, the resulting output pulses provide a measure of the magnitude and sense of the frequency error. This frequency steering ultimately allows the loop to correct the frequency error until the phase error is within the PFD linear range. Unlike the steady-state condition of a second-order PLL with an integrator loop filter, the final phase error in the $\Delta\Sigma$ FD will never reach zero. This is solely due to the quantizer continually providing an incorrect measure of the second integrator output so the actual phase error is never corrected.

Characterizing the acquisition phase can be done by simulating the non-linear $\Delta\Sigma$ FD model with some initial state and observing if it acquires lock. There are two scenarios to consider — initial acquisition after applying power and potential re-acquisition after a large input frequency step. The first case implies that the second integrator has an initial state of zero and the PFD will have an arbitrary initial phase error. From that point on, the discriminator should begin to cycle slip and eventually acquire lock. This is clearly seen in Figure 4.18 where the two integrator states are plotted over samples of the reference

Figure 4.18: $\Delta\Sigma$ FD initial acquisition after power-up.

period T_r . The initial phase error oscillates between positive and negative limits during cycle slipping while the second integrator exhibits damped transient oscillation indicating convergence toward steady-state phase error. During this period, the 1-bit quantizer output produces limit cycles with progressively shorter spans. After acquiring lock, the phase error brackets the zero degree phase point, causing the second integrator output never to asymptotically reach zero. At this time, the $\Delta\Sigma$ FD is functioning normally by producing a random output due to the noise shaping.

The second case assumes that the $\Delta\Sigma$ FD is initially in steady-state mode and is then subjected to a large input frequency step within its input dynamic range. In this case, the question is whether the discriminator loses lock at all, and if so, does it reacquire lock?

Figure 4.19: $\Delta\Sigma$ FD acquisition following an input frequency step.

Figure 4.19 shows an example of steady state operation near midband followed by a large frequency step such that the final state is near the operating limits of the discriminator. Results of this simulation show that the loop remains locked (i.e. $|\phi_{err}| < 2\pi$) so acquisition is not an issue. The smooth transition between the two operating points can easily be seen by observing the output bitstream. Initially, the density of zeros and ones is approximately equal indicating midband operation. After the input frequency step, the density of ones is much higher than the zero density which shows a shift in operation toward the upper limits of the discriminator input range. If the shift is too large, the discriminator output would saturate to a steady stream of ones (or zeros for a downward shift) because its dynamic range has been exceeded.

4.1.5 Achievable Signal-to-Noise Ratio

The achievable signal-to-noise ratio (SNR) of a circuit depends on its architecture and the quality of the implementation. In this section, the emphasis is on identifying what architectural parameters alter the theoretical peak SNR of the second-order $\Delta\Sigma$ FD. Any non-ideal effects due to circuit implementation are discussed in the following section from an abstract perspective.

The discriminator architecture thus far is a single-loop structure employing single-bit quantization. Arguments in favor of this choice were discussed earlier, but it is nevertheless useful to know whether a significant improvement in performance can be obtained through a small change in the discriminator architecture. In some cases, the parameter choice is restricted by the intended application so no change is permitted (e.g. input frequency deviation determined by modulation scheme). If that is the case, it imposes a minimal level of performance on the proposed discriminator architecture including effects of non-ideal implementation.

Two parameters that are solely defined by the intended application are the input frequency deviation and the signal BW. Recall that the input deviation in a $\Delta\Sigma$ FD is analogous to the AC level of conventional $\Delta\Sigma$ modulators. Thus the achievable SNR depends upon the ratio of the input frequency deviation to the reference frequency in the signal bandwidth of interest. In this architecture, the maximum input dynamic range is equal to the reference frequency so the maximum input deviation is restricted to

$$f_{dev(max)} = \pm \frac{f_r}{2} \quad (4.19)$$

if the carrier frequency is set exactly at midband. Since the $\Delta\Sigma$ FD modulus n is controlled directly by the modulation signal, the operating point is always at midband. Operating at midband makes full use of the available dynamic range and results in the best quantization noise shaping due to the high quantizer gain K_q . This is an advantage over conventional $\Delta\Sigma$ modulators which may exhibit poor noise shaping when the input signal is near the dynamic range limits. Although f_r defines the theoretical dynamic range, the usable

range is somewhat less than that. For input signals with a very small frequency deviation, the discriminator noise floor will overpower the detected signal resulting in <0 dB SNR. Conversely, if the frequency deviation is too large, the discriminator will begin to overload and the in-band noise floor will increase. The overload conditions will also cause in-band spurious tones to be present and both of these effects will degrade the SNR as shown in Figure 4.20. The SNR improves by 6dB/octave from the noise floor until overload occurs

Figure 4.20: Signal-to-noise ratio of second-order frequency discriminator (BW=200KHz).

where it quickly decreases.

For a fixed input deviation, the achievable SNR is directly related to the signal BW. That is to say, the amount of quantization noise (albeit shaped) will increase for larger frequency offsets from the carrier. Since the signal BW is imposed by the application, the goal is to either reduce the total amount of quantization noise and/or shift any residual in-band noise elsewhere.

Although the input signal deviation and BW are fixed, there is some freedom in

choosing the sampling frequency of the $\Delta\Sigma$ FD. Since the modulator reference frequency f_r is identical to the $\Delta\Sigma$ FD sampling frequency, it is desirable to use an integer multiple of the data rate, which for GSM is 270.83Kbits/s. This simplifies the design of the transmit filter in the modulation path. The incoming RF signal to the discriminator must also be oversampled to shift enough of the quantization noise out of band. Considering these two factors, a minimum sampling frequency of 13MHz was chosen, with other valid choices being $f_r = n \cdot 13\text{MHz}$ where $n \in \{1, 2, 3 \dots\}$. The oversampling ratio (OSR) is defined as

$$\text{OSR} = \frac{f_r}{2\text{BW}} \quad (4.20)$$

The improvement in SNR for a second-order $\Delta\Sigma$ modulator is approximately 15dB/octave of oversampling [Cand85] and in the context of the discriminator, the same relationship applies for a fixed input signal with one exception. As the reference (sampling) frequency increases, the input dynamic range increases proportionally. Thus for a fixed input

Figure 4.21: Effect of oversampling ratio on SNR of second-order frequency discriminator with BW=200KHz.

frequency deviation, the discriminator effectively sees a progressively smaller signal as f_r increases (e.g. doubling f_r results in an *effective* 6dB attenuation of the input signal). The net improvement in SNR, with increasing oversampling ratio reduces to 9dB/octave as illustrated in Figure 4.21.

Improvement of the overall discriminator SNR can also be obtained by using more values to represent the signal, which is realized by increasing the quantizer resolution. Previously, it was stated that a 1-bit (two level) quantizer was chosen to reduce the multi-modulus divider complexity (see Table 4.1) and also eliminate any non-linearity due to threshold error in the quantizer. If hardware complexity were not an issue, the expected SNR improvement for using a multi-level (fractional resolution) quantizer are illustrated

Figure 4.22: Effect of quantizer resolution on SNR of second-order frequency discriminator with BW=200KHz.

in Figure 4.22. It is clear that the SNR improves by approximately 5dB for each doubling of the number of quantizer levels.

The final architectural change one may contemplate is to design a multi-modulus divider that can divide by fractional increments. An expression for the available moduli in

such a divider would be

$$\text{modulus} \in (n, n + k, n + 2k, \dots) \quad k < 1 \quad (4.21)$$

where n is an integer and k is typically a rational fraction. This effectively reduces the minimum frequency step of the divider output to a fraction of the reference frequency f_r . Similar to increasing the number of quantizer levels, dividing by a fractional value reduces the frequency error of the divider output with respect to the reference frequency. This reduction of quantization error results in a 6dB/octave net improvement of the discriminator SNR as seen in Figure 4.23. Bear in mind the design of the fractional divider

Figure 4.23: Effect of divider fractional- δ on SNR of second-order frequency discriminator with BW=200KHz.

becomes increasingly difficult for smaller fractional divisions since the divider must resolve a fraction of the input frequency period. For example, if the goal is to design a fractional divider with moduli $(n, n + 0.25, n + 0.5, \dots)$ and the RF input frequency is 2GHz, the divider must be able to accurately resolve 125ps increments. This time period is close to the propagation delays of fast digital gates so it is difficult to make such an accurate measurement.

4.1.6 Influence of Circuit Parameters

Previously, the emphasis was to explore the single-loop discriminator in Figure 4.1 from an architectural perspective, while continually assuming an ideal implementation could be realized. While this approach helps in the initial analysis, it is obvious that such a realization is impossible and one must account for non-ideal circuit parameters. Even though the final circuit is unknown, it is possible to determine the effect of some parameters by modelling each functional block as before and including the non-ideal effects. The following non-ideal effects are considered:

- phase-frequency detector deadzone
- charge pump leakage and mismatch
- quantizer threshold offset

Note that circuit noise is not considered here because the various noise spectral densities can only be extracted from the final circuit realization, so this analysis is left for Section 4.2.6.

Deadzone in a phase-frequency detector (PFD) refers to a region of operation where the PFD produces no output (except noise) even though there is an input signal present. In digital PFD's, the deadzone exists when the input phase error is small (e.g. locked condition for a PLL). The PFD is unable to produce the corresponding narrow UP and DN pulses and instead produces no output which implies an effective PFD gain $K_\phi = 0$. This can be a serious problem since operating in the deadzone opens the discriminator loop until the random phase error grows large enough to produce a finite PFD output signal. The net effect on the overall discriminator SNR is shown in Figure 4.24 where an increase in the deadzone destroys the quality of the noise shaping in-band. This in turn raises the amount of in-band noise while the signal remains constant so the SNR is reduced.

Figure 4.24: Effect of PFD deadzone on SNR of second-order frequency discriminator with BW=200KHz.

The task of the second integrator in the forward path is to integrate the phase error represented by the PFD output ($UP - DN$). This function is best realized using a charge pump (CP) which adds charge to or removes charge from a capacitor depending on

Figure 4.25: Simplified single-ended charge pump.

whether the phase error is positive or negative. Ideally the CP output voltage is

$$v_o(t) = K_{CP} \int [v_{UP}(t) - v_{DN}(t)] dt \quad (4.22)$$

where K_{CP} is the CP gain and the input is the time difference between the *UP* and *DN* pulses. The equivalent frequency domain representation can be expressed as an ideal integrator

$$H_{CP}(s) = K_{CP} \left(\frac{1}{s} \right) \quad (4.23)$$

Any leakage component from various sources can be lumped into one equivalent resistance R in parallel with the CP capacitor C_p as indicated in Figure 4.25. This modifies the original transfer function from that in Equation (4.23) to

$$H_{CP}(s) = K_{CP} \left(\frac{1}{s + \frac{1}{RC}} \right) \quad (4.24)$$

which is equivalent to a lowpass filter with a pole at $\omega = 1/(RC)$. This has a direct consequence on the $\Delta\Sigma$ FD performance because the inverse noise transfer function which originally was two cascaded integrators is now an integrator and a lowpass filter. The quality of the noise shaping will suffer depending on how far the CP pole has shifted from the ideal $\omega = 0$ position due to leakage.

Mismatch in the CP current sources is another potential hazard which effectively causes the CP to exhibit a different gain K_{CP} depending on whether the phase error is positive or negative. In conventional PLL's this causes the reference frequency to feed through to the output which is undesirable. A similar effect would occur in this discriminator architecture if multi-level quantization is used. The argument here is that multi-level quantizers have a well defined gain and any non-linear magnitude error at its input is passed through to its output. However, if only a single-bit quantizer is used, the quantized output is simply the *sign* of the input voltage with respect to some reference level. This effectively shields any magnitude error due to current source mismatch from the discriminator output.

Threshold errors in a multi-level quantizer result in non-linear conversion gain from the input to the output even if the overall response is monotonic. This is undesirable in any $\Delta\Sigma$ modulator and efforts should be made to minimize it. A two level (i.e. 1-bit) quantizer has only one threshold the input voltage is compared to. Does the position of this threshold really matter as in the multi-level quantizer case? The answer depends on the type of noise shaping filter chosen. More importantly, it is the filter DC gain that matters, since it lies in the forward path of the discriminator loop and an offset error is really a DC error. If there is sufficient forward gain (e.g. double integrators), then the threshold offset error can be compensated for by a shift in operating point with no adverse effect on performance.

4.2 BiCMOS $\Delta\Sigma$ Frequency Discriminator Chip

The central component in this new modulator architecture is the $\Delta\Sigma$ frequency discriminator. The performance of the discriminator has significant impact on the overall modulator performance so a great deal of effort has been invested into its design. The goal was to realize the entire discriminator, whose architecture is depicted in Figure 4.1, in monolithic form for the following reasons:

- an integrated chip is more cost effective than a discrete design
- significant power savings can be achieved through integration

The performance requirements of the $\Delta\Sigma$ FD are unlike conventional analog-to-digital (A/D) converters that operate at lower IF frequencies, since it must function at the RF carrier frequency without loss of performance. This requirement alone imposes a challenge to devise an architecture that can operate at high speed without the need to use exotic integrated circuit technologies as in [Jens95]. Since the target standard is DCS-1800 (or DCS-1900), the $\Delta\Sigma$ FD should have a usable bandwidth greater than 2GHz. This precludes the use of current standard CMOS processes for the high speed blocks, so bipolar devices were chosen instead. With careful design techniques, a circuit realized in a bipolar process can achieve a lower power consumption than a comparable CMOS solution. Some of the $\Delta\Sigma$ FD blocks use CMOS devices to improve their performance, so

using a BiCMOS technology results in a single-chip design. A 3V power supply voltage was used to further control the overall power consumption and ensure compatibility with battery operated handsets.

The $\Delta\Sigma$ FD functional specifications, listed in Table 4.2, were derived from the GSM modulation standard and those necessary for modulation. The upper RF frequency is

Table 4.2: $\Delta\Sigma$ FD functional specification for GSM modulation.

RF Range (MHz)	Reference f_r (MHz)	Modulus n (bits)	Modulation a (bits)
500-2500	40	6	3

chosen for operation beyond 2GHz, but in essence the lower range can extend far below 500MHz, limited only by the range of moduli in the divider. The maximum reference frequency is 40MHz for wideband modulation schemes but for the GSM modulator a reference of $f_r = 13\text{MHz}$ was used. The modulus range n of the divider defines the RF input frequency range for a given reference frequency. Using a 6-bit word enables one of 64 different moduli to be chosen. Modulation in the form of coarsely dithered bits is generated by the Gaussian filtered and equalized data symbols. The dynamic range of this signal determines the required modulation input range of the $\Delta\Sigma$ FD. Note that these specifications exceed that what is necessary for the DCS-1800 standard which allows the $\Delta\Sigma$ FD chip to be used in other applications as well.

A discussion of high-speed low-power circuit design and process techniques will be presented first. These techniques form a basic design strategy that maximizes the speed potential of the BiCMOS devices for a given power consumption. The entire power budget is further controlled by employing power saving schemes wherever possible. What follows is detailed circuit descriptions of each block along with simulated results.

4.2.1 High Speed, Low Power Design Techniques

High operating speed is achieved by using fully differential ECL/CML logic with a small voltage swing wherever possible. This provides a high immunity to common-mode noise

while the small signal swing reduces the switching time for a given slew rate. Such a design strategy will satisfy the speed requirements, but the resulting power consumption may be compromised since traditional ECL/CML logic circuits were implemented as discrete functions with 50Ω output drivers. The high currents (and hence high power) necessary to drive the 50Ω transmission lines, would seem to render this approach unsuitable.

The entire discriminator is realized in monolithic form as a single BiCMOS chip, so the internal load impedances can be much higher, thus reducing the required drive current for each logic block. Although this reduces the power consumption dramatically, further improvements can still be made. Unlike in CMOS circuits, where the dynamic power consumption varies with operating frequency, the maximum frequency of a bipolar circuit depends on the current density through each device [Roul90]. The design strategy is to supply just enough current density through each device such that it will operate at the

Figure 4.26: Transit frequency of a 1X bipolar device ($A_E = 0.8 \times 4.0 \mu m$) [Hada91].

desired speed, not its *maximum* speed. The relationship between operating speed and current density is found through examination of the bipolar devices transit frequency f_T , which is related to the base transit time, or time it takes for carriers to travel through the base. Typically, the speed of a device will peak near a point known as the critical current density J_K , which is near the peak f_T , as shown in Figure 4.26. Operating a device at this point achieves the highest speed for a given technology and as the current density is reduced, the maximum speed decreases. Increasing the current density beyond the critical level reduces the f_T of the device, so this offers no speed advantage.

This characteristic can be exploited in the multi-modulus divider architecture, since only the dual-modulus divider operates at the RF frequency, while the counter and decoder both operate at a reduced speed (i.e. by at least a factor of 4 in this implementation). Therefore, a reduction in power consumption can be realized by independently adjusting the current densities of each block. What may not be obvious at this point, is how to adjust each device's current density, when only a single supply current is available. To understand this, a closer examination of a typical ECL/CML logic gate depicted in Figure

Figure 4.27: A typical ECL/CML logic gate.

4.27 is necessary. The current source for the logic transistors sets the absolute current which is steered through one of the paths depending on the inputs. Control of the current density J_K through each device is accomplished by adjusting its emitter area A_E such that

$$A_E = \frac{I_{TAIL}}{J_K} \quad (4.25)$$

Clearly, the tail current must be set high enough to ensure a sufficient current density for a minimum size device (e.g. typically denoted as a 1X device) or maximum speed cannot be achieved. For technologies where only a fixed set of device sizes are available (as is the case with the BiCMOS process used here), the nearest size that guarantee the desired operating speed is chosen. Customizing the emitter area is desired for the logic core, but output loading (i.e. fan out) may dictate a higher tail current to minimize propagation delay. With knowledge of the output loads of each stage, the driver currents can be optimized to reduce the power consumption while maintaining the desired operating speed. An interesting characteristic of differential ECL/CML circuits is that there is always a constant tail current flow, regardless of the operating frequency. Logic is realized by current *steering*, not current *switching* as is the case in CMOS logic. This implies that even if there is no logic activity, the power consumption remains constant whereby the dynamic power consumption of CMOS logic varies directly with frequency (i.e. $P_D \propto C_L f (V_{DD})^2$). One method of overcoming this limitation is to switch off the current sources during periods of inactivity, which of course is application specific. This is easily accomplished by removing the bias voltage in each current source which disables it so only leakage current flows. Note that output driver stages could use pull-down resistors as opposed to using an active pull-down to gain some advantage in switching speed. If this is the case, the tail current cannot be switched off for these output stages.

Further exploitation of each tail current is possible by realizing more logic functions in each gate. This has an immediate effect on power consumption, since fewer gates are required and combining some gates into a single complex gate reduces the total propagation delay. More logic functionality can be implemented in one gate by introducing additional logic signal levels that are separated by DC offsets as shown in

Figure 4.28: Complex ECL/CML logic gate.

Figure 4.28. The limit to the number of levels is determined by the available voltage headroom dictated by the supply voltage. Each logic transistor must be biased to prevent deep saturation (i.e. $V_{CE} \approx V_{BE}$) to allow it to operate at the desired speed. Further restrictions are imposed by the load resistors where the output voltage is developed and the tail current sources which must operate in their compliance region. In this design, three logic levels are possible with a 3V supply voltage with only mild saturation at supply voltages down to 2.7V (i.e. minimum voltage of a 3V battery).

4.2.2 Multi-Modulus Divider with Low Delay

This section describes a new low-delay multi-modulus divider architecture suitable for use in the single-loop second-order $\Delta\Sigma$ frequency discriminator. The requirement for low delay arises from the single-loop discriminator model in Chapter 3, which was derived from the multi-loop second-order structure (see Figure 3.10). Manipulation of the blocks into the single-loop structure is only possible if the multi-modulus divider does not

introduce additional delay (i.e. a z^{-1} term anywhere in the path). This implies that the modulus of the divider can be altered and take effect *before* the current cycle is completed. The new multi-modulus divider architecture shown below is composed of an input buffer,

Figure 4.29: Block diagram of low delay multi-modulus divider.

a 4/5 dual-modulus divider (DMD), a 6-bit programmable synchronous counter clocked by the DMD, and a decoder. A wide range of moduli are possible by selectively controlling the modulus of the dual modulus divider through detection of different states programmed via the A and B registers in the decoder. Earlier multi-modulus architectures reported in [Perr97] and [Fili97] cannot be used in the single-loop discriminator since they introduce additional delay.

Operation of the multi-modulus divider begins by pre-loading a value into the synchronous M -counter and initially setting the dual-modulus divider to divide by its lower modulus $n = 4$. The A and B control inputs to the decoder, select two states within the range M that toggle the DMD modulus. The state diagram for a complete cycle

Figure 4.30: State diagram of low delay multi-modulus divider.

is illustrated in Figure 4.30. From this state diagram, an expression for the overall divider modulus is

$$N = 4M + 5 + A + B \quad M \geq 12 \quad (4.26)$$

where

M = coarse channel select

A = modulation control

B = feedback control

Thus the supported modulus range for a 6-bit M -word, a 3-bit A -word and a 2-bit B -word is

$$N \in \{53, 54, \dots, 264\}$$

which provides ample range for many synthesizer applications.

The multi-modulus divider can be partitioned into a high-speed section including the input buffer and DMD operating at the RF frequency. The synchronous counter and decoder blocks are clocked by the DMD output which has a maximum frequency of $1/4$ the RF input frequency. This is the exact situation where power savings can be made by designing the blocks to operate at their highest required frequency rather than the speed the technology can offer.

The RF buffer serves two functions — it provides an impedance-matched interface between the VCO signal and the divider and it shifts the DC point to a level which is

Figure 4.31: RF buffer with level shifter.

Figure 4.32: RF input buffer gain and bandwidth.

compatible with the clock input of the divider. The bandwidth of this amplifier must exceed the desired 2GHz operating frequency derived from the DCS-1800 specification. A schematic of the fully differential buffer is shown in Figure 4.31. The buffer contains an on-chip input impedance matching network to couple to the VCO, a DC level shifting section and on-chip bias networks. Note that the RF input signal is AC coupled into the buffer so on-chip biasing sets the DC level before the emitter followers such that their output is compatible with the differential logic switches. The logic tail current is mirrored from a reference whose current is programmable. A DC level shift is introduced by pulling a constant tail current through a load resistor. The AC performance of this buffer is characterized in Figure 4.32 which shows it has a BW extending beyond 2GHz.

The dual-modulus divider uses a synchronous, fully differential architecture to provide maximum speed with high common-mode noise rejection. This block is clocked by the RF buffer output and therefore operates at the RF frequency. For this reason, it will also consume the most power (per function). The modulus choice arises from a compromise between lowering the clock speed of the low-speed section at the expense of a more complicated DMD architecture. Choosing a DMD modulus of 4/5, allows the M -counter and decoder to operate four time slower with a minor penalty in DMD power consumption. The DMD is designed around three flip-flops with multiple feedback paths as shown in Figure 4.33. Modulus control is accomplished by selecting the appropriate state to inject into the last flip-flop with a 2-input MUX. Note that there are only three gate delays in the DMD since the NAND gate and MUX are absorbed into the flip-flops using ECL/CML stacked logic. While this reduces the power consumption, more importantly, it enhances the maximum speed of the divider by minimizing the total signal path length. The operation of the DMD is illustrated in the timing diagram of Figure 4.34 for a 2GHz input frequency.

The remaining blocks are the M -counter and the decoder, both which are clocked by the DMD output. Both these blocks are realized using similar ECL/CML techniques as in the DMD except the devices are scaled for a lower current density since their maximum speed is $f_{in}/4$. This produces a favorable power reduction for these blocks without any performance penalty.

Figure 4.33: Differential 4/5 dual-modulus divider.

Figure 4.34: Dual-modulus divider timing diagram.

Figure 4.35: Multi-modulus divider timing diagram.

Timing for the complete multi-modulus divider during locked conditions is portrayed in Figure 4.35. This simulation illustrates the low-delay behavior of the multi-modulus divider through a modulus change during the current divide cycle. The sequence of events for the various annotated time points are as follows:

- 1 - start a new divide cycle by pre-loading the M -counter (Q_4 - Q_0)
- 2 - change the modulus by altering B_0 during current cycle on rising edge of f_r
- 3 - detect the new value of B and switch DMD to $\div 5$ mode
- 4 - detect the value of A and switch DMD back to $\div 4$ mode
- 5 - output pulse when M -counter reaches terminal count=0

Earlier it was mentioned that an import requirement of this multi-modulus divider architecture is that it does not introduce additional delay in the discriminator loop. What this implies is that the divider modulus must be able to be changed (updated) *after* a divide cycle has begun. Most divider architectures [Perr97],[Fili97] respond to a modulus change request at the beginning of the *following* cycle, which isn't acceptable in this $\Delta\Sigma$ FD architecture. A key point in this divider design is that the modulus setting affects the end of the current divide cycle and not the beginning. Therefore, it is permissible to begin the divide cycle and subsequently change the modulus as long as the setup time is met. This restriction is shown in Figure 4.36 which focuses near the end of the divide cycle. The sequence of events in Figure 4.36 at the various time points are:

- 1 - change the modulus by altering B_0 during current cycle on rising edge of f_r
- 2 - detect the new value of B and switch DMD to $\div 5$ mode
- 3 - detect the value of A and switch DMD back to $\div 4$ mode
- 4 - output pulse when M -counter reaches terminal count=0
- 5 - compare divider phase on falling edge of f_r

The actual setup time is the point where the decoder detects the M -counter state, so its A and B inputs must be stable at that time. From the overall multi-modulus timing diagram

Figure 4.36: Multi-modulus divider setup time.

in Figure 4.35, the A and B inputs change on the rising edge of f_r , while the multi-modulus divider terminal count is the falling edge of f_r (i.e. one half of a reference period). Now the worst setup condition occurs when $A = 7$ (since it's inverted) and $B = 3$ so that 12 DMD clock cycles are required while it's in $\div 5$ mode. If no setup violation is to occur, the time for 12 DMD cycles is restricted to

$$\frac{T_r}{2} > 12 \cdot T_{DMD} \quad (4.27)$$

Since the DMD is in $\div 5$ mode (worst case), this implies the RF input period must be

$$\frac{T_r}{2} > 12(5 \cdot T_{in}) \quad (4.28)$$

yielding an equivalent RF frequency

$$f_{in} > 120 \cdot f_r \quad (4.29)$$

This restriction limits the ratio of RF input and reference frequencies that may be used in the final modulator. Previously, a 13MHz reference frequency was chosen which is suitable for the DCS-1800 standard which operates in the 1.8GHz frequency band.

4.2.3 Phase-Frequency Detector

Initial acquisition and phase error detection are performed in the phase-frequency detector (PFD). During acquisition mode, there is a frequency error between the reference frequency f_r and the multi-modulus divider output. Reduction of the acquisition time can be achieved by employing some form of frequency steering. After the $\Delta\Sigma$ FD loop is locked, the PFD must provide an accurate measure of the phase error once per reference period.

The digital PFD architecture in Figure 4.37 was chosen to provide initial frequency steering and also produce bipolar phase error signals (i.e. UP and DN), which readily interface to the analog charge pump. This architecture doesn't exhibit any asynchronous race problems (due to logic feedback) found in some popular commercial parts [Moto83], which eliminates any glitches from occurring at the PFD outputs. This is an important benefit, since the charge pump continually monitors the PFD outputs and reacts to any signal activity including inadvertent glitches. The method used to prevent race conditions is to use an edge triggered flip-flop with an asynchronous reset, whose schematic is shown in Figure 4.38.

Figure 4.37: Phase-frequency detector with asynchronous reset.

The flip-flop consists of a master and slave latch clocked by opposite phases of the clock. Normally a logic one is latched and transferred to the output after a clock edge. If reset is active, the latched value is overloaded regardless of the clock state which forces it to the reset state $Q = 0$.

Reference and divider edges are captured by pre-loading a logic one at the flip-flop inputs and allowing the signal edge to clock the data value to the output. Depending on which edge arrives first, either the *UP* or *DN* output is activated. The other delayed input signal edge subsequently activates the remaining PFD output. At this point in time, both

Figure 4.38: Differential PFD flip-flop with asynchronous reset.

Figure 4.39: Phase-frequency detector timing diagram.

outputs are active which is clearly shown in the timing diagram of Figure 4.39. The feedback NOR gate asynchronously resets the flip-flops after some delay (typically the gate propagation delay). Note that by extending the feedback path delay, the reset signal can be arbitrarily delayed. This feature is used in the PFD design to effectively stretch the width of the *UP* and *DN* pulses. This is done to prevent the PFD from trying to produce infinitely narrow pulses when the phase error is small (e.g. during locked conditions). Since the charge pump is measuring the difference between *UP* and *DN*, it doesn't matter whether they are both at a logic one or zero, so no error is introduced. If the PFD were forced to resolve a very small phase error, it would be unable to produce the

corresponding narrow pulse width due to the logic finite rise and fall times. This region is known as the PFD deadzone where the gain K_ϕ drops to zero. Operating the discriminator in this region is analogous to opening the loop where phase control is lost, so it is important to reduce, or ideally eliminate the deadzone. If the PFD pulse shapes have identical rise and fall times, delaying the reset would completely eliminate the deadzone, but in practice there may be some mismatch. Simulation results for the transfer characteristics of this PFD are shown in Figure 4.40 where the deadzone has been

Figure 4.40: Differential phase-frequency detector transfer function.

effectively eliminated. There is a fixed offset of 2mV at $\phi_{err} = 0$ (not visible in the plot) but this will simply cause a fixed phase offset in the discriminator loop, which has no adverse effect on performance.

4.2.4 Charge Pump

Double integration of the frequency error is performed first by the implicit conversion of frequency-into-phase in the PFD, and again in the charge pump. The first integration is inherently ideal, but the second, implemented as an analog charge pump, is subjected to various non-ideal effects. Proper design of the charge pump is imperative to retain the desired quantization noise shaping. The function of the charge pump is to integrate the differential phase error pulses from the PFD and produce an analog voltage that is subsequently quantized.

Figure 4.41 shows a simplified example of a differential charge pump being driven by the PFD *UP* and *DN* signals. The source or sink currents are selectively steered (no

Figure 4.41: Simplified representation of a differential charge pump.

current switching occurs) to charge and discharge the capacitor so that the differential output voltage is

$$v_o(t) = K_{CP} \int [v_{UP}(t) - v_{DN}(t)] dt \quad (4.30)$$

where K_{CP} is the CP gain and the input is the difference between the *UP* and *DN* pulses.

The two output signals from the PFD represent four possible states, leading to four modes of operation for the CP as illustrated in Table 4.3. The performance of the charge pump

Table 4.3: Differential charge pump operating modes.

Mode	PFD outputs		Active sinks	
	UP	DN	COUT	$\overline{\text{COUT}}$
off	0	0	$\overline{\text{UP}}$	$\overline{\text{DN}}$
discharge	0	1	$\overline{\text{UP}}$, DN	none
charge	1	0	none	UP, $\overline{\text{DN}}$
reset	1	1	DN	UP

will be compromised by several factors:

- PFD deadzone distorts phase error pulse shape
- mismatch in the source and sink currents will introduce non-linear effects

The PFD deadzone issue was addressed earlier and through careful design of the PFD, it was not visible in simulation results.

Mismatch in the source and sink currents effectively changes the integration time constant I/C for positive and negative phase errors. This affects charge pumps that use current steering and current switching during periods where the PFD outputs are active. However, when the PFD outputs are inactive in a switched CP, the current sources are turned off, with no further change of the CP output assuming leakage is not an issue. This is not the case in a current steered CP, since the sources are never turned off but simply rerouted. The worst situations occur during modes (off) and (reset) (see Table 4.3). The (reset) mode is not as critical since it has a short duration, but the (off) mode is active for all times other than the charge or discharge times. During the (off) mode, a *net* current of zero should exist leaving the capacitor voltage unchanged, but mismatch will generate a finite differential current causing the capacitor voltage to drift.

Preventing current source mismatch is inherently a difficult task since deviation in the process parameters can alter the active device characteristics. Although there are circuit techniques that are less sensitive to process changes, there will always be some residual

effects present. Rather than attempting to use static circuit techniques, an active feedback network is used to control the current sources in the charge pump shown in Figure 4.42.

Figure 4.42: Simplified schematic of differential charge pump with active feedback.

The charge pump architecture is fully differential and makes use of current steering rather than switching. This method avoids switching the upper PMOS current sources which are much slower than the NPN current sinks. Current is either directed into or diverted from the holding capacitor, depending on the state of the PFD output. Since both current sources are always on, they must be matched, or the capacitor voltage will drift over time. The feedback circuit monitors the common-mode voltage of a replica of the actual charge pump core circuit and adjusts the upper PMOS sources to maintain that common-mode voltage. Success of this approach depends on the matching of the replica and the core circuit which amounts to matching of the V_{BE} and V_T of the bipolar and MOS devices respectively. The sensitivity to V_{BE} mismatch in the lower current sources is reduced by emitter degeneration, while careful layout techniques ensure minimal mismatch for the PMOS current sources. Simulation results demonstrating the CP operating modes with a

Figure 4.43: Differential charge pump timing diagram.

sampling period of 50ns are shown in Figure 4.43. The effect of any current source mismatch that may still exist is somewhat reduced by having a smaller sampling period. This reduces the time spent in the (off) mode assuming the average phase error remains constant.

The allowable DC range of the charge pump refers to the extent the capacitor voltage can change, without the upper or lower current sources operating outside of their compliance voltage. The lower limit is loosely determined by the levels of the *UP* and *DN* logic signals who set the point where the NPN differential current switches begin to saturate. The upper limit is set by the capacitor voltage that causes the lower PMOS transistors to shift from saturation into linear region of operation. Using a cascode configuration for the PMOS current sources reduces the available voltage range, but it raises the current source output impedance, improving the overall linearity. The charge pump linearity and voltage range are plotted in Figure 4.44. From this plot, it is clear that the common-mode voltage should be set midway between the upper and lower limits to maximize the usable range. For added flexibility, this common-mode voltage is externally

Figure 4.44: Differential charge pump linearity and output range.

adjustable. A worst case current mismatch of 0.8% occurs over the ± 1 V range. Since the nominal steady-state phase error is small in locked conditions, the CP output voltage will hover around the common-mode voltage V_{CM} where there is far less mismatch.

4.2.5 Quantizer

1-bit quantization of the differential charge pump output voltage is realized with a differential comparator. The charge pump voltage is compared to a reference voltage and the quantizer output reflects whether the input is above or below the reference voltage, after the clock edge. The advantage of a 1-bit (2-level) comparator is that it is inherently linear and a misplaced reference level simply introduces an offset that is quickly corrected by the high forward gain (i.e. double integrators) of the $\Delta\Sigma$ FD. For multi-level quantizers, reference threshold errors are somewhat more serious, since they introduce non-linearities in its gain while the gain of a 1-bit quantizer is undefined. In conventional $\Delta\Sigma$ modulators,

a multi-level quantizer is accompanied by a corresponding D/A converter in the feedback path. The linearity of the D/A converter ultimately limits the linearity of the modulator since D/A converter non-linearities are directly added to the signal. This limitation does not apply to the single-loop $\Delta\Sigma$ FD since the feedback signal is entirely digital.

The DC gain of the quantizer must be high enough to resolve small differential charge

Figure 4.45: Differential 1-bit quantizer with BiCMOS input buffer/comparator.

pump voltages while not interfering with the operation of the charge pump itself. This implies that the small amount of charge stored on the integrating capacitors must not be affected by the quantizer input stage, so a very high input impedance is necessary. The quantizer configuration consists of a high gain BiCMOS buffer, followed by a latched comparator as illustrated in Figure 4.45. The BiCMOS buffer consists of an NMOS differential gain stage followed by a cascade of NMOS/bipolar differential gain stages. The overall DC gain of the buffer is 36dB over a 400MHz bandwidth. The high input impedance of the NMOS first stage prevents loading of the integrating capacitors and the remaining stages increase the signal amplitude while providing the necessary DC levels for the comparator.

The DC sweep in Figure 4.46 of the quantizer input buffer shows an input sensitivity

Figure 4.46: Differential 1-bit quantizer DC transfer characteristic.

of $\pm 2.5\text{mV}$ which ensures full output voltage swing when the differential CP voltage is small. The high gain is necessary to ensure the latch has a sufficient input signal level to prevent metastability, which would degrade the performance of the $\Delta\Sigma\text{FD}$.

4.2.6 Noise Calculations

Chapter 3 introduced the linear model for the second-order single-loop frequency discriminator in Figure 4.47 which served two purposes:

- reduce the non-linear elements into their quasi-linear equivalents
- introduce various lumped noise sources that model equivalent circuit noise

Now that the actual circuit design is known, the various noise sources can be quantified and their effect on the discriminator performance can be determined.

Figure 4.47: Second-order single-loop $\Delta\Sigma$ frequency discriminator model.

Using this model, the output of the $\Delta\Sigma$ FD was expressed earlier as

$$f_{out} = \frac{N}{f_{vo}} f_v - n + (1 - z^{-1})^2 S_Q + \frac{N f_v}{2\pi f_{vo}} (1 - z^{-1}) S_{DIV} + \frac{N f_v}{2\pi K_\phi f_{vo}} (1 - z^{-1}) S_{PFD} + \frac{N f_v}{2\pi K_{CP} K_\phi f_{vo}} (1 - z^{-1})^2 S_{CP} \quad (4.31)$$

where the lumped noise sources are defined as:

$$\begin{aligned} S_Q &= \text{quantizer noise} \\ S_{DIV} &= \text{divider noise} \\ S_{PFD} &= \text{phase detector noise} \\ S_{CP} &= \text{charge pump noise} \end{aligned}$$

Quantization noise is modeled using the arguments in [Benn48], which maps the non-linear quantization process into an ideal conversion with added noise. The remaining noise sources are actual $\Delta\Sigma$ FD circuit noise effects and can be calculated by performing Hspice[®] noise analysis on the chip implementation. These results are then mapped into the form expected by the model (i.e. V^2/Hz , rad^2/Hz etc.).

Quantization noise generated by the $\Delta\Sigma$ FD was described earlier as a white noise source that is shaped by an equivalent highpass filter of the same order as the $\Delta\Sigma$ FD. The justification for modeling the quantizer as an additive white noise source is based on work by Bennet [Benn48], who replaced a deterministic non-linearity with a stochastic linear system, thereby permitting the use of linear system analysis in a non-linear system containing a quantizer. This linear approximation is based on the following properties of the quantization error:

- uncorrelated with input signal
- uniformly distributed over the quantization interval
- uniform power spectral density

Although quantization noise does not have these properties, the additive white noise approximation is applicable for higher order $\Delta\Sigma$ FD's, provided the quantizers do not overload [Cand85]. Assuming the above properties are valid, the uncorrelated noise power (variance) of a quantizer with step size Δ is defined as

$$e_{rms}^2 = \frac{\Delta^2}{12} \quad (4.32)$$

which is uniformly distributed over the range $(-\Delta/2, \Delta/2)$. In the context of this $\Delta\Sigma$ FD, a unity step in the quantizer is equivalent to an input frequency step of f_r (i.e. the $\Delta\Sigma$ FD

output represents a frequency). However, the quantization noise of the 1-bit quantizer in the $\Delta\Sigma$ FD model (see Figure 4.47) is output referred so $\Delta = 1$. Given that a quantized signal sampled at f_r has all its noise power aliased into the frequency band $(-f_r/2 \leq f \leq f_r/2)$, the spectral density of the $\Delta\Sigma$ FD quantizer noise is

$$S_Q(f) = \frac{1}{12f_r} \quad [\text{Hz}^2/\text{Hz}] \quad (4.33)$$

This white noise is shaped through feedback so the power spectral density of the frequency quantization noise at the output of the $\Delta\Sigma$ FD becomes

$$S_{Qout}(f) = |(1 - e^{-j2\pi f T_r})^2|^2 \cdot \frac{1}{12f_r} \quad [\text{Hz}^2/\text{Hz}] \quad (4.34)$$

Divider phase noise is a direct result of small variations in the transition times (i.e. timing jitter) of the divider output. Intuitively, for a multi-stage divider architecture, the total jitter at the output is related to the jitter in previous stages. This statement implies that asynchronous dividers inherently have a higher phase noise than synchronous dividers, since each stage independently contributes a noise component (i.e. it is uncorrelated) [McC192]. The approach used in obtaining the total noise at the multi-modulus divider output is to calculate the total output-referred noise voltage density when the divider is biased at the switching point (i.e its metastable point). In a simulation environment, this is accomplished by forcing the differential clock signal levels such that $V_{CLK} = \overline{V_{CLK}}$. One may argue that the noise level (and distribution) is signal dependent, but the equivalent phase noise is only significant during the switching time. That is to say, any noise voltage

Figure 4.48: First-order mapping of voltage noise to timing jitter.

generated by the divider has little effect on the timing jitter (and phase) except during the switching time as illustrated in Figure 4.48. The effect of noise is to shift the threshold crossing time by an amount proportional to the voltage error divided by the slew rate. Using this approximation, the timing jitter can be written as

$$\Delta t = \frac{\Delta v_n}{\text{slew rate}} \quad [\text{s}] \quad (4.35)$$

From Equation (4.35), the timing jitter variance in the divider output is

$$\overline{\Delta t^2} = \frac{\overline{\Delta v_n^2}}{(\text{slew rate})^2} \quad [\text{s}^2] \quad (4.36)$$

Converting the timing variance to phase noise with respect to the reference period yields

$$S_{DIV} = \left(\frac{2\pi}{T_r}\right)^2 \cdot \overline{\Delta t^2} = \left(\frac{2\pi}{T_r \cdot \text{slew rate}}\right)^2 \cdot \overline{\Delta v_n^2} \quad [\text{rad}^2] \quad (4.37)$$

Similarly, the phase noise spectral density with units $[\text{rad}^2/\text{Hz}]$ can be written in the same form as Equation (4.37) assuming the noise voltage is expressed in $[\text{V}^2/\text{Hz}]$. The equivalent output-referred divider noise is found by applying the appropriate transfer function giving

$$S_{DIVout}(f) = \left(\frac{Nf_v}{2\pi f_{vo}}\right)^2 |1 - e^{-j2\pi f T_r}|^2 \cdot S_{DIV}(f) \quad [\text{Hz}^2/\text{Hz}] \quad (4.38)$$

The $(1 - e^{-j2\pi f T_r})$ term indicates the divider noise is differentiated by the action of the $\Delta\Sigma\text{FD}$ loop. This is a desirable characteristic since it shifts some of the phase noise power out of band, unlike in conventional fractional-N synthesizers. Inspection of the transfer function seen by the divider noise reveals that it is modulated (i.e. a signal dependent gain) by the RF input signal f_v . This turns out not to be an issue of concern if actual values for the gain parameters are used. Since narrowband FM modulation is used with a carrier $f_{vo} > 1\text{GHz}$, the deviation of f_v from the nominal carrier frequency f_{vo} is minimal, so $f_v(t) \approx f_{vo}$. This effectively sets the gain of the divider and PFD phase noise to a fixed value.

Noise originating in the PFD manifests itself as timing jitter in the edges of the UP and DN pulses. However, unlike the divider noise, the PFD timing jitter is converted into an equivalent noise voltage seen at its output. Note that the width of the differential phase error pulse ($UP - DN$) is the true output, so the timing jitter on both rising *and* falling edges contribute to the total PFD noise as shown in Figure 4.49. Conversion of the timing

Figure 4.49: Differential PFD output timing jitter.

jitter into an equivalent output noise voltage is done by finding the average jitter over one reference period, converting to an equivalent voltage using slew rate and multiplying by the PFD gain which gives

$$S_{PFD} = \left(\frac{2\pi K_{\phi}}{T_r \cdot \text{slew rate}} \right)^2 \cdot \overline{2\Delta v_n^2} \quad [\text{V}^2] \quad (4.39)$$

Similar to the divider noise density, the PFD noise density can be output referred to get

$$S_{PFDout}(f) = \left(\frac{Nf_v}{2\pi K_{\phi} f_{vo}} \right)^2 |1 - e^{-j2\pi f T_r}|^2 \cdot S_{PFD}(f) \quad [\text{Hz}^2/\text{Hz}] \quad (4.40)$$

Noise in the charge pump appears at its output as random fluctuations of the capacitor voltage due to passive and active devices. This noise is dominated by the current sources (bipolar and CMOS) that charge and discharge the CP capacitors. The amount of noise generated in the CP also depends on its mode of operation (i.e. signal dependent) so the analysis uses the reset mode which represents the worst case conditions where both current sources are active. Practically, the reset mode has a short duration, so the actual average noise will be lower than predicted.

Results from the noise analysis can be directly used in the $\Delta\Sigma\text{FD}$ model since they are

computed as a noise voltage density. Referring the CP noise to the $\Delta\Sigma$ FD output gives

$$S_{CPout}(f) = \left(\frac{Nf_v}{2\pi K_{CP} K_\phi f_{vo}} \right)^2 |(1 - e^{-j2\pi f T_r})|^2 \cdot S_{CP}(f) \quad [\text{Hz}^2/\text{Hz}] \quad (4.41)$$

indicating the output referred CP noise is double differentiated which reduces the in-band noise.

Figure 4.50 shows the output-referred contributions of all noise sources as well as the total noise. These plots indicate that the $\Delta\Sigma$ FD in-band noise is dominated by the PFD up

Figure 4.50: $\Delta\Sigma$ FD output referred frequency noise spectral density.

to 10 KHz and then the second-order shaped quantization noise begins to dominate. The quantization noise is lowpass filtered within the main modulator loop but the in-band $\Delta\Sigma$ FD noise is potentially more serious. This noise imposes a limit on the smallest resolvable frequency that can be detected within a desired bandwidth.

A key parameter when designing a demodulator for FM signals is the minimum discernible frequency or phase that can be detected. In the GMSK modulator, the $\Delta\Sigma$ FD in the feedback path acts as a demodulator of the VCO signal. Obviously, the demodulator architecture imposes some ideal lower detectable limit, but the circuit noise in the actual implementation will degrade this value.

Results from the previous noise analysis can be used to calculate the discriminator frequency sensitivity by referring the total noise to the input. Input-referred noise is simply the output-referred noise divided by the signal transfer function which in this case is

$$\begin{aligned} S_{fi}(f) &= \left| \frac{1}{H_{SIG}(f)} \right|^2 \cdot S_{fo}(f) & [\text{Hz}^2/\text{Hz}] \\ &= (f_r)^2 \cdot S_{fo}(f) \end{aligned} \quad (4.42)$$

Once the noise is input referred, it can be compared to a signal of equal power (i.e. 0dB SNR). The frequency deviation of this input signal represents the minimum detectable frequency that is resolvable by the $\Delta\Sigma$ FD.

Although the second-order discriminator is designed to demodulate an FM signal, it can also detect a PM signal by integrating the digital output frequency to get the phase modulation

$$\phi_o(z) = \frac{2\pi}{f_r(1-z^{-1})} \cdot f_o(z) \quad (4.43)$$

Bear in mind that the original second-order frequency noise shaping (i.e. 40dB/dec) will be reduced to first-order phase noise shaping as a result of this integration. This infers that the $\Delta\Sigma$ FD will exhibit better performance when used to demodulate an FM signal rather than a PM signal.

The input phase sensitivity can be found using the same approach used to calculate the input frequency sensitivity, and recalling that phase is simply the integral of frequency. With this relationship, the input-referred phase noise spectral density is

$$S_{\phi_i}(f) = \left| \frac{2\pi}{f_r(1 - e^{-j2\pi f T_r})} \right|^2 \cdot S_{f_i}(f) \quad [\text{rad}^2/\text{Hz}] \quad (4.44)$$

It is usually desired to know the phase noise in terms of the actual single sideband (SSB) power with respect to the carrier level expressed as

$$\mathcal{L}(f) = \frac{\text{sideband power density}}{\text{carrier power}} \quad [\text{dBc/Hz}] \quad (4.45)$$

The phase noise spectral density will be equivalent to the actual side-bands if the peak phase fluctuations are much less than one radian. In modulation theory, this is equivalent to saying that the high-order modulation components are insignificant compared to the fundamental modulating frequency. This narrow band assumption holds for the GMSK modulator because the GSM modulation frequency is much less than the carrier frequency. From small angle modulation theory, the SSB phase noise is

$$\mathcal{L}(f) = \frac{S_{\phi}(f)}{2} \quad [\text{dBc/Hz}] \quad (4.46)$$

where the units [dBc/Hz] represent the ratio of sideband power in a 1Hz BW at some frequency offset to the carrier power.

Tables 4.4 and 4.5 compare the simulated minimum discernible frequency and phase for the ideal and transistor level $\Delta\Sigma$ FD models. The maximum SNR is obtained by exercising the full dynamic range of the discriminator by setting the peak deviation to $f_{dev} = \pm f_r/2$. The results are somewhat optimistic since a linear discriminator model

Table 4.4: Ideal $\Delta\Sigma$ FD input sensitivity.

Signal BW (KHz)	Frequency ($f_r = 13\text{MHz}$)		Phase
	SNR (dB)	f_{min} (Hz)	ϕ_{min} (deg)
25	112.8	11	0.03
50	97.8	60	0.09
100	82.8	334	0.25
200	67.8	1881	0.71
400	52.8	10581	1.99

Table 4.5: Optimal $\Delta\Sigma$ FD input sensitivity.

Signal BW (KHz)	Frequency ($f_r = 13\text{MHz}$)		Phase
	SNR (dB)	f_{min} (Hz)	ϕ_{min} (deg)
25	107.4	20	0.07
50	95.7	75	0.13
100	82.1	359	0.28
200	67.6	1917	0.73
400	52.8	10632	2.01

was used which doesn't predict any non-linear effects such as quantizer overload and in-band spurious tones. The ideal and transistor level performance of the discriminator is very similar for bandwidths exceeding 50KHz, where the quantization noise dominates above any circuit induced noise. Below this frequency, the PFD noise dominates degrading the overall SNR and achievable minimum detectable frequency and phase.

4.2.7 Mixed Signal Design and Layout Techniques

Electrical interference between the discriminator analog and digital circuit blocks and their sensitivity to noise are important issues to consider when designing such a large mixed-signal integrated circuit. In this case, the problem is further compounded by frequencies ranging from DC to several GHz coexisting on the same silicon substrate. Noise coupling between circuits can occur:

- through the substrate
- via the power supplies
- due to electro-magnetic radiation

Design strategies to circumvent any potential problems can be applied at the circuit and physical levels. The primary goal is to reduce the noise generated on-chip such that any residual noise will not adversely affect the performance of other circuits. Decoupling the circuit blocks from each other using both electrical and physical means is the next step to control the effect of any residual noise. Finally, architectural techniques can be applied to

reduce the sensitivity of a circuit to any external noise that cannot be avoided. A clever design strategy will employ all of these techniques to control the effects of noise.

Digital switching noise is a common problem in mixed-signal IC's which degrades the performance of analog blocks elsewhere on the substrate. Switching noise results from changing current flow while a logic gate switches states. The fast surge of current causes a voltage drop in the power supply due to the inductance of the bond wires and package pins [Rain94]. The disturbance of the supply voltage can affect any other circuit powered by the same supply. Historically, discrete designs capacitively decoupled the power supply at each load but on IC's the required capacitance would consume a large area leading to higher silicon costs. An alternative approach is to reduce (or even eliminate) any switching currents by using fully differential structures that *steer* rather than *switch* current. The advantage is the elimination of any current changes, since a constant current flows through the tail of each differential pair. ECL and CML circuits are classic examples of differential circuits and are used throughout the $\Delta\Sigma$ FD implementation. Single-ended signals (i.e. differential logic with a single output) may be used to restrict the total power consumption for areas with minimal logic activity such as static inputs etc. The rationale is that a constant current flow through a differential output doesn't generate noise but simply wastes power, so it's better to remove the complementary output. Any residual switching current effect can be minimized by ensuring the power supply lines have low impedance. The common V_{CC} plane in the core of Figure 4.51 forms the low impedance supply path for the dual-modulus divider while the ground (V_{EE}) is a ring around the entire circuit. It is also important to ensure that supply paths converge at one point (i.e. a star configuration) which helps to isolate local currents flowing in one branch from affecting another.

One or more sources injecting noise directly into the substrate causes currents to circulate, subsequently contaminating other circuits. In silicon technologies, the substrate doping results in a poor but finite conductivity as opposed to pure silicon which is an insulator. The substrate acts like a large backplane, so once a noise current enters the

Figure 4.51: Layout plot of high-speed differential 4/5 dual-modulus divider.

substrate, it can flow into other circuits affecting their performance. Isolation between circuits is accomplished by exploiting features of the technology (e.g. trench isolation) and/or maintaining electrical separation. While physical separation can be used in lightly doped substrates, the low resistance of heavily doped substrates makes this an ineffective way to reduce coupling [Su93]. Electrical separation can be realized through a reverse biased PN junction or guard rings both which interrupt the noise current and divert it elsewhere. The reversed biased junction forms a depletion capacitor physically sandwiched between the substrate and the circuit. Any noise currents that couple through this capacitor are shorted to the opposing plate which is connected to V_{CC} or V_{EE} (i.e. AC ground). Reverse biased junctions are used to shield capacitors and PMOS devices from

substrate noise. Guard rings offer an effective way of controlling substrate noise currents by providing a barrier around a circuit which is connected to a low impedance return path [Su93]. The key to effective noise control is to bias the guard ring at the appropriate point. This implies that separate guard paths are required that terminate off-chip. Figure 4.51 shows the guard ring around the dual-modulus divider which ultimately is terminated at the appropriate point (not shown).

Improving a circuit's noise immunity provides an effective method of tolerating any residual noise that could not be avoided, or to simply provide a higher noise margin. Signal integrity must be maintained everywhere, so noise immunity is important not only for each circuit block but also along the interconnects. The transfer of a signal from one point to another can be impeded by noise along the path and reference variations due to the lack of a common reference node (i.e. ground is not always a suitable reference node). The use of differential signals can mitigate both problems since any noise along the path is common mode which is rejected and the local reference node is embodied within the signal itself. If a signal is measured relative to a reference, it is important that this reference be on-chip rather than off-chip. The reason for this is that an external reference rarely matches its on-chip counterpart due to package parasitics etc. Furthermore, the ideal reference should track the absolute signal deviation due to power supply voltage and temperature drift. This approach is used throughout the discriminator design which greatly reduces its sensitivity to power supply variation and thermal drift. Reduction of circuit sensitivity to substrate noise can be improved by ensuring that the noise appears as common-mode to a signal. Differential circuits are quite effective in this respect, since noise on the ground lines is inherently common-mode and is rejected [Maki95]. Power supply noise is not suppressed in the same manner, so effort must be made to keep this noise minimal. Physical symmetry is exploited to keep substrate noise at common levels. This is especially important in the differential analog charge pump capacitors which carry little charge, so they are laid out in a symmetrical fashion as seen in Figure 4.52. Although

Figure 4.52: Layout plot of differential BiCMOS charge pump.

noise will accumulate or deplete charge from the capacitors altering their absolute voltage. the differential signal remains intact so performance is not compromised.

4.2.8 Measured Results

The $\Delta\Sigma$ FD chip is partitioned into the blocks shown in the photomicrograph of Figure 4.53. Due to the mixed signal architecture, strategic floor planning is used (in addition to other techniques) to minimize signal interference between blocks. The discriminator was fabricated in a standard $0.8\mu\text{m}$ BiCMOS process [Hada91] and covers an area of 9mm^2 . The total power consumption is 75mW with a 3V power supply.

Preliminary testing of the BiCMOS $\Delta\Sigma$ FD verified the circuit functionality and static

Figure 4.53: Photomicrograph of BiCMOS $\Delta\Sigma$ frequency discriminator.

DC performance. The DC test results listed in Table 4.6 show that the majority of the

Table 4.6: BiCMOS $\Delta\Sigma$ FD chip DC test results.

Module	Current (mA)	Power ($V_{CC}=3V$) (mW)	Power (%)
Divider	19	57	76
Phase detector	2	6	8
Charge pump	3	9	12
Quantizer	0.4	1.2	1.6
Logic	0.4	1.2	1.6
Reference	0.2	0.6	0.8
Total	25	75	100

power consumption is attributed to the multi-modulus divider. This is expected, since it operates at the RF frequency, which forces the devices to operate with a much higher current density. In fact, newer bipolar technologies (e.g. SiGe) would offer much lower current consumption for the same operating frequency since they have a higher transit frequency f_T .

Further testing of the chip was done to measure the limits of its dynamic performance and are listed in Table 4.7. These results show that the single-chip discriminator can operate at the desired DCS-1800 frequency band (i.e. 1900MHz). The input sensitivity is -10dBm or equivalently an amplitude of $\pm 100\text{mV}$, which is the minimum level required to toggle the RF input buffer. Reliable operation of the discriminator at this input level and

Table 4.7: BiCMOS $\Delta\Sigma$ FD chip AC test results.

Maximum RF (GHz)	Maximum f_r (MHz)	Input level (dBm)	Sensitivity (dBm)
2.5	50	-2	-10

frequency is not guaranteed due to the sinusoidal shape of the input signal, which may cause the RF buffer (see Figure 4.31) to revert to its metastable state. The ideal signal shape is one with steep edges to reliably toggle the input stage of the RF buffer. However, at 2GHz, the harmonics necessary to reconstruct an approximate square wave (i.e. 1st, 3rd,

5th etc.) would require a much larger BW. Assuming a carrier input signal $A \sin(\omega_c t)$, the switching speed is defined by the slope or derivative

$$\frac{d}{dt} A \sin(2\pi f_c t) = A 2\pi f_c \cos(2\pi f_c t) \quad (4.47)$$

From Equation (4.47), the input amplitude A and the carrier frequency f_c directly influence the maximum switching speed. Since the carrier frequency is assumed to be relatively constant, increasing the signal amplitude will improve the switching speed, limited only by the maximum input level the RF buffer can tolerate. Typically, the normal input level should be around -2dBm, which is sufficient to ensure correct operation of the $\Delta\Sigma$ FD. In the context of the GMSK modulator, these input requirements are easily met by providing some gain after the VCO to compensate for any signal attenuation or low VCO output power.

Spectral analysis of the $\Delta\Sigma$ FD output bitstream will reveal its actual noise shaping action under various input conditions. There are two general cases to consider — a DC input consisting of an unmodulated carrier and an AC input represented by an angle modulated carrier. A DC input is used to force the discriminator to produce limit cycles, leading to discrete spurs in the output spectrum. Note that this is analogous to injecting a DC voltage or current into a conventional $\Delta\Sigma$ modulator. The length of the limit cycles depends on the discriminator order and the actual DC input level. This mode of operation also exercises very little of the available dynamic range (i.e. the input is not changing), leading to the worst possible quantization noise shaping. Another way of stating this is that the presence of spurs is an impairment to the ideal quantization noise shaping defined by the discriminator order. A common method of overcoming DC limit cycles is to introduce an AC component to the input signal. The trivial case occurs when the input signal itself is AC (e.g. an FM modulated carrier) or it can result from some intentional dither added to an existing DC input signal. The effect the AC signal has on the discriminator is to randomize its internal states such that it exhibits chaotic behaviour, which results in optimal quantization noise shaping.

Measured and simulated results are compared in Figure 4.54 for a 1.8525GHz carrier

frequency. Comparing the two curves shows that the hardware discriminator suffers from a higher in-band noise floor than expected. The exact reason for the poor noise shaping was partly due to the phase noise of the $\Delta\Sigma$ FD reference, but may also be due to its internal phase-frequency detector, whose noise dominates in-band (refer to Figure 4.50). An exact source for the additional noise was not determined, although reference and PFD noise affect the discriminator in the same way. Since the discriminator lies in the feedback path of the GMSK modulator, its noise is not attenuated at the output. This will directly impact the overall phase noise floor of the GMSK modulator, as will be discussed in Section 5.5.2.

Figure 4.54: Output spectrum of $\Delta\Sigma$ FD with DC input (unmodulated carrier).

The second case to consider is the noise shaping action of the $\Delta\Sigma$ FD with an angle modulated input. Injecting an FM signal into the frequency discriminator eliminates any periodicity in the output bitstream since it adds an AC component to the DC carrier

Figure 4.55: Output spectrum of measured $\Delta\Sigma$ FD bitstream with 100KHz single-tone FM modulated carrier.

frequency. This has the benefit of forcing the discriminator internal states to remain randomized, which eliminates any limit cycles that cause spurs. A measured output spectrum with an input signal comprised of a 1.8525GHz carrier modulated by a 100KHz single-tone FM signal is shown in Figure 4.55, with an in-band view in Figure 4.56. In the in-band view, the spur at 100KHz represents the original FM modulating tone that has been demodulated by the discriminator (recall a $\Delta\Sigma$ FD is an FM demodulator *and* A/D converter). The height of this spur is directly proportional to the original frequency deviation of the transmitted signal. Any non-linear characteristic of the discriminator will produce distortion in the output spectrum, which is evident by the presence of the 200KHz spur (i.e. at twice the modulating frequency).

Figure 4.56: In-band view of measured $\Delta\Sigma$ FD output spectrum with 100KHz single-tone FM modulated carrier.

This chapter presented a new $\Delta\Sigma$ frequency discriminator architecture and monolithic implementation that is suitable for use in a GMSK modulator. Experimental measurements showed that the RF bandwidth of the discriminator was 2.5GHz, which covers the entire DCS-1800 frequency range. It was further revealed that the integrated $\Delta\Sigma$ FD has a higher in-band noise floor than expected from the simulation results, which will impact the performance of the GMSK modulator. Further indirect testing of the $\Delta\Sigma$ FD, embedded inside the GMSK modulator loop, is covered in Chapter 5.

Chapter 5

Modulator Design and Implementation

Chapter 3 presented a new wideband GMSK modulator architecture using a $\Delta\Sigma$ frequency discriminator based synthesizer. Although this architecture is suitable for any wideband, continuous-phase modulation scheme, emphasis will be placed on the GSM standard which belongs to the class of GMSK modulation schemes. The following section defines the necessary synthesizer parameters that are used to realize the GSM modulator.

5.1 GSM Design Example

The GSM standard is a member of a class of spectrally efficient modulation schemes that use Gaussian filtering of the data to yield an RF spectrum that has most of its power contained within a desired bandwidth (i.e. 200KHz for GSM). It uses a data symbol rate of 270.833Kb/s and a maximum frequency deviation of ± 67.71 KHz dictated by the minimum frequency shift keying requirements. There are presently two standards that use this form of modulation; GSM which uses the handset frequency band 890-915MHz and DCS-1800 whose handset frequency band is 1710-1785MHz [GSM98]. Since the trend in future wireless radio systems is an increase in carrier frequency, the DCS-1800 frequency band will be used to illustrate the feasibility of this architecture for high carrier frequency transmitter designs. A number of reasonable assumptions will be made for various components of the modulator to determine a final parameter set without knowledge of

Figure 5.1: Block diagram of the $\Delta\Sigma$ FD based GMSK modulator.

actual circuit implementation details (described in Section 5.2 and 5.3). Figure 5.1 illustrates the complete modulator block diagram that will be used to compute the necessary design parameters. The synthesizer will be restricted to second order ($n = 2$) to relax the loop stability requirements. The digital loop filter is chosen to be an integrator with phase lead compensation with a transfer function

$$F(z) = a_2 + \frac{a_3}{1 - z^{-1}}$$

where a_2 and a_3 are the proportional and integral gains respectively. Additional filtering of the $\Delta\Sigma$ quantization noise outside of the loop bandwidth (as yet undefined) is realized with a lowpass digital filter. Both digital filters are realized in the DSP block in Figure 5.1. The remaining forward path of the main synthesizer loop consists of a D/A converter, analog integrator and VCO all of which are assumed ideal with some conversion gain K . The strategy used to determine the loop parameters takes advantage of the fact that in this architecture, the synthesizer noise performance and modulation capability can be set independent of each other. The goal then is to adjust the loop parameters to meet the noise specification and subsequently design the modulation path transfer function to provide the

necessary Gaussian filtering and compensate for the closed-loop response of the synthesizer. The foremost parameter that influences the choice of other parameters is the reference frequency f_r , which is also the sampling frequency for the DSP. Since the modulator architecture contains many discrete-time blocks, f_r must be set high enough for a given loop bandwidth to reduce the effect of aliasing to acceptable levels. A more restrictive lower limit is set by the modulation bandwidth, which by definition is much greater than the synthesizer open-loop bandwidth, and loosely defines the dynamic range requirements. The final restriction in choosing a reference frequency arises from the need to have an integral number of samples per data symbol to simplify the digital transmit filter design. Taking these issues into account, a reference frequency of 13MHz is chosen which satisfies:

$$f_r \gg \text{open-loop BW}$$

$$\text{samples/symbol} = \frac{13\text{MHz}}{270.833\text{Kb/s}} = 48$$

Phase noise in most $\Delta\Sigma$ fractional-N type synthesizers is largely a result of the quantization noise of the $\Delta\Sigma$ modulator ($\Delta\Sigma$ frequency discriminator in this case). The key to controlling its noise contribution is to attenuate it before it appears at the VCO output. This is performed implicitly by the closed-loop transfer function of the synthesizer which acts as a lowpass filter. The bandwidth of the synthesizer closed-loop response is determined by the open-loop bandwidth BW which also governs the transient response of the synthesizer. There is usually a design trade-off when choosing the open-loop bandwidth since a narrow BW is desired for noise performance, while a wide BW is necessary to improve switching speed. One way of circumventing this dilemma is to widen the open-loop bandwidth dynamically during switching, and reduce it after acquiring lock by modifying the loop filter parameters. This approach has been used in previous analog PLL designs at the risk of potentially introducing switching transients. In the architecture of Figure 5.1, on the other hand, the entire loop filter is realized in the digital domain, so changing the loop filter dynamics amounts to varying a filter coefficient and any switching transients are avoided. Since the loop parameter requirements during

acquisition may now be ignored, a suitable open-loop bandwidth for adequate noise performance is 30KHz using a 13MHz reference frequency (determined through simulation). The remaining synthesizer parameter to adjust is the natural frequency ω_n (or equivalently the damping factor ζ) which controls the stability of the loop. From Section 3.3, the equations governing these parameters are

$$\begin{aligned} BW &= K'(a_2 + a_3) \quad [\text{rad/s}] \\ \omega_n &= \sqrt{-K' f_r \ln \left[\left(\frac{a_2}{a_2 + a_3} \right)^{(a_2 + a_3)} \right]} \quad [\text{rad/s}] \\ \zeta &= \frac{\sqrt{K'(a_2 + a_3)}}{2 \sqrt{-f_r \ln \left(\frac{a_2}{a_2 + a_3} \right)}} \end{aligned}$$

where a_2 and a_3 are the proportional and integral gains of the digital loop filter, and K' is the open-loop gain not including the loop filter gain. The choice of a_2 and a_3 is restricted to integral values to permit the use of fixed point arithmetic in the DSP, while K' can be any reasonable value since it is partly defined by the analog components in the synthesizer loop. Choosing $a_2 = 512$ and $a_3 = 1$ results in the synthesizer loop parameters shown in Table 5.1.

Table 5.1: Synthesizer loop parameters for GSM modulation.

f_r	BW	ω_n	ζ
13MHz	30KHz	11KHz	1.4

The meaning of the damping factor ζ in this architecture is analogous to that in a continuous-time PLL, but sampling in the discrete-time blocks degrades the available phase margin for frequencies approaching the reference frequency f_r . Outside the open-loop bandwidth, additional poles are placed to further filter the quantization noise. These are realized in DSP as a digital Butterworth lowpass filter of order $n = 4$ with a pole frequency $f_p \approx 100\text{KHz}$. Using these parameters, the ideal synthesizer open-loop

Figure 5.2: Open-loop transfer function of $\Delta\Sigma$ FD based synthesizer.

response is shown in Figure 5.2.

Once the synthesizer noise performance has been met, design of the modulation data path can proceed. The goal is to synthesize a digital transmit filter that combines the Gaussian pulse shaping and the equalizer transfer functions. The Gaussian filter response is solely defined by the chosen modulation standard and is obtained by sampling the impulse response of an ideal Gaussian filter convolved with a rectangular pulse spanning the symbol period T_b . This is necessary because it is more practical to inject the data symbols as impulses at intervals of T_b (i.e. ± 1 data symbols injected at the symbol rate) rather than at the oversampled rate f_r . For a normalized Gaussian bandwidth $BT = 0.3$, there is significant ISI from the adjacent symbol on either side of the current symbol. Thus the filter impulse response, which theoretically has infinite duration, can be truncated to span at least three symbol periods and appropriately windowed to reduce the effect of truncation. The resulting tap weights must be quantized in a manner that retains the

desired filter response.

Design of the equalizer is simplified by taking advantage of the fact that only a limited modulation bandwidth needs to be equalized, since most of the signal energy at higher frequencies has been attenuated by the Gaussian filter. The closed-loop response of the synthesizer can be computed using the models developed in Chapter 3 along with the parameters defined for the desired application. Once the Gaussian filter and equalizer responses are known, an equivalent FIR filter can be designed using a filter synthesis algorithm similar to the method described in Section 3.2.2 on model development.

5.2 Mixed-Signal Synthesizer Blocks

This section discusses the implementation strategies of the remaining main synthesizer components:

- digital signal processor
- digital-to-analog converter
- analog integrator
- voltage-controlled oscillator.

Theoretical analysis and design implementation issues for the $\Delta\Sigma$ frequency discriminator chip were described earlier in Chapter 4.

These components, along with the $\Delta\Sigma$ FD form the main synthesizer loop. The final modulator prototype combines the $\Delta\Sigma$ FD chip and several discrete components on a custom printed circuit board (PCB).

5.2.1 Digital Signal Processor

The fixed-point digital signal processor (DSP) is realized using a field-programmable gate arrays (FPGA), rather than a custom application specific integrated circuit (ASIC). Mapping the design into an FPGA allows various architectures to be quickly evaluated without incurring the cost and delays of IC fabrication. The design flow in Figure 5.3

shows the transformation of generic VHDL (very large scale hardware description language) code into a routed FPGA. Logic synthesis and compilation was completed using

Figure 5.3: FPGA design flow.

the Synopsys[®] synthesis tools to produce an optimized gate netlist. The Xilinx[®] FPGA vendor tools were used to complete the mapping and routing of the final design into an FPGA.

The DSP is partitioned into two major functional blocks — a $\Delta\Sigma$ quantization noise filter and a synthesizer loop filter. The noise filter is used to attenuate the $\Delta\Sigma$ quantization noise (in addition to the synthesizer closed-loop response) originating from the $\Delta\Sigma$ FD and the digital $\Delta\Sigma$ modulator in the modulation data path. It also indirectly controls the amount of data equalization required depending on the rate of attenuation. The loop filter combines the desired synthesizer loop filter response along with dynamic range control and interface logic to interface to the external digital-to-analog (D/A) converter. Since both filters are in the main synthesizer loop, it is imperative to minimize the latency through the DSP to ensure overall loop stability. In this implementation, there is a single clock cycle delay due to each digital filter which was taken into account in the stability analysis in Section 3.3.3.

The quantization noise filter has a Butterworth lowpass response that has a flat pass-band gain which introduces no magnitude distortion to the modulation data (up to the first spectral lobe) and thereafter has a monotonic attenuation which relaxes the design of the equalizer. There are a number of issues to consider when choosing the filter order and pole frequency. First, it is imperative to set the pole frequency outside of the synthesizer loop BW. This will ensure the quantization noise filter does not interfere with the synthesizer

loop stability, which should predominantly be defined by the digital loop filter. The second, more subtle point, relates to how the response of the noise filter affects the corresponding equalizer response. Recall that the purpose of modulation data equalization is to compensate for attenuation due to the synthesizer loop BW since the modulation BW is much higher. Increasing the modulation BW implies that more aggressive equalization is required since the higher signal frequencies suffer more attenuation. Conversely, a narrower modulation bandwidth requires much less equalization, assuming a constant loop BW. There is an optimal amount of equalization that should be used to guarantee that a sufficient amount of the $\Delta\Sigma$ FD dynamic range is exercised. A comparison of the 270.833Kb/s GSM modulation data rate to the 13MHz $\Delta\Sigma$ FD dynamic range (equal to the reference f_r) shows that the data occupies a small percentage of the available dynamic range prior to equalization. Now the 30KHz loop BW will require some equalization, but by introducing even more closed-loop attenuation with the quantization noise filter, the required amount of equalization can be increased to an optimal level. It is difficult to calculate the necessary filter response based on the desired maximum equalized data amplitude, since there is no single closed form solution. This problem can be circumvented by performing time-domain simulation of the GMSK modulator with various filter responses and noting the maximum amplitude of the resulting equalized data. An acceptable solution is one that yields the desired equalized data amplitude using the least complex quantization filter response. Using this approach, a Butterworth filter with the parameters in Table 5.2 satisfies the $\Delta\Sigma$ FD dynamic range requirements.

Table 5.2: Digital Butterworth filter parameters.

Sampling rate (MHz)	filter order	f_{pole} (KHz)
13	4	100

Realization of the digital filter in Table 5.2 presents a number of problems. A Butterworth filter has an infinite impulse response (IIR) which implies it is a recursive structure with poles and zeros. Realizing such a filter using fixed-point arithmetic is difficult for two reasons — the filter coefficients must be quantized into integers, which

alters the pole/zero locations, and the filter states will suffer from quantization effects due to the feedback paths. The first reason, quantizing the filter coefficients, is more of an issue for the poles, especially if they are near $z = \pm 1$ on the unit circle. The sensitivity of the poles in this region can result in a final filter response dramatically different from the desired response. Even if one were able to find a set of quantized coefficients that produced an acceptable response, the signal quantization of internal nodes would preclude the use of a simple IIR structure. This is not a unique problem and many methods to realize IIR filters have been reported [John93],[Demp98],[Laak98]. Rather than contending with realizing an IIR filter, a finite impulse response (FIR) filter is used. The consequence of this is that it takes a much higher FIR filter order to approximate the original Butterworth filter response. However, there are several fixed-point implementation benefits of an FIR filter:

- no multipliers are necessary if the partial sums of the filter taps are pre-computed and stored in a ROM
- the maximum dynamic range of each internal node can be calculated so signal quantization effects can be completely avoided

A partial block diagram of the FIR filter structure is shown in Figure 5.4(b) where the input data addresses the contents of a bank of ROMs. The outputs of the ROMs contain the

Figure 5.4: Mapping the (a) direct form FIR filter into (b) a ROM based FIR filter.

partial sums for a sequence of concurrent taps of the filter. These partial sums are summed together to form the final filter output. The filter implementation is further simplified by shifting its location ahead of the loop filter in the DSP block. In doing so, a 1-bit wide addressing scheme can be used, controlled by the 1-bit $\Delta\Sigma$ FD output bitstream. This removes the need for any address decoding other than a delay line consisting of a chain of flip-flops. Note that the addition in Figure 5.5(a) is done in parallel to enhance the maximum filter speed through a reduction in the number of adder levels. A decision must be made in choosing the number of ROMs. Using more ROMs results in a smaller ROM

Figure 5.5: Equivalent filter structures: (a) smaller ROM size with more adder levels or (b) larger ROM size with less adder levels.

size but more partial sums are produced, increasing the number of adders. Conversely, using fewer ROMs implies less addition, but each ROM must contain more partial sums that have a larger range of values so a wider data word size is required. This concept is illustrated in Figure 5.5 for a simple 8-tap FIR filter. Using the same approach, the memory and adder cost of the 256-tap FIR filter used in this implementation is summarized in Table 5.3 for various ROM sizes including a direct implementation (i.e. no

Table 5.3: 256 tap FIR filter partitioning.

# ROMs	address size (bits)	ROM size (bits)	total memory (bits)	# adders	adder depth (levels)
none	—	—	—	255	8
128	2	7x4	3.5K	127	7
64	4	8x16	8K	63	6
32	8	9x256	72K	31	5
16	16	10x65536	10M	15	4

ROMs). From Table 5.3, it is clear that there is a trade-off between computational effort (addition) and total memory requirements. A direct implementation has a high computational penalty, while large ROM sizes produce fast results at a high memory cost. In this implementation, 64 ROMs are used to improve the maximum filter speed while using only 8 K of memory.

Mapping the original Butterworth filter response into an FIR structure begins by computing the ideal impulse response sampled at f_r , as in Figure 5.6(a). The length of the FIR filter (number of taps) depends on the original IIR impulse response length. Theoretically, the IIR impulse response is infinitely long, but one can truncate (and optionally window) the response to yield the approximate finite impulse response in Figure 5.6(b). In this case, the resultant FIR impulse response is truncated to 256 samples and windowed to reduce the end effects. The samples of the finite impulse response are the exact multiplier coefficients required in the FIR implementation. However, for fixed-point arithmetic, the multiplier coefficients must be converted into integers, as in the IIR case.

Figure 5.6: Quantization noise filter: (a) ideal infinite impulse response and (b) scaled and quantized finite impulse response.

There are many ways to quantize the coefficients, ranging from simple truncation and rounding, to passing the coefficients through a non-linear system to obtain a new set with different characteristics. An example of the latter case is to $\Delta\Sigma$ modulate the high resolution impulse response into a finite set of integers [Rile94]. The quantization error introduced is spectrally shaped to shift the error to higher frequencies so the in-band filter response is retained. Since the purpose of this filter is to attenuate $\Delta\Sigma$ quantization noise generated by the $\Delta\Sigma$ FD and digital $\Delta\Sigma$ modulator in the modulation path, additional $\Delta\Sigma$ noise shaping serves no purpose. Simple rounding along with dynamic range scaling can be used to convert the floating point coefficients into integers as shown in Figure 5.6(b). Scaling the filter gain is done to make full use of the available dynamic range in the ROMs, which results in less error between the full precision and quantized filter response. Frequency domain verification of the final filter, in Figure 5.7, shows that the FIR filter in-

Figure 5.7: Frequency response of ideal Butterworth IIR filter and approximate FIR filter.

band response has not been altered much, but the stop-band attenuation has been reduced to 40dB compared to the original IIR filter. This amount of stop-band filtering is enough to adequately attenuate the quantization noise from the $\Delta\Sigma$ FD and digital $\Delta\Sigma$ modulator.

Synthesizer loop parameters are controlled in the digital loop filter which forms the second section of the DSP. The loop parameters were evaluated in Section 5.1 to provide adequate noise suppression. The loop filter in Figure 5.8, comprised of an integrator with phase lead compensation, is realized as a digital accumulator with a feedforward path. Both the proportional and integral gains are adjustable to even powers of two (i.e. no multipliers required) to alter the loop dynamics in real time. This feature accounts for the different loop parameters required during acquisition as opposed to noise suppression (i.e. a wide loop BW during acquisition and a narrow loop BW for noise suppression). Signal scaling, in conjunction with multi-bit buses, is used to control the dynamic range of each

Figure 5.8: Digital synthesizer loop filter employing saturation arithmetic and detection.

node. This minimizes the probability of limit cycles due to signal overflow, which would result in extremely non-linear behavior. While signal scaling can reduce the probability of overflow, it may not eliminate it. For example, the initial state of the loop filter after power is applied is usually not controllable, although steps have been taken in this design to ensure the filter resets to a known state. A more serious condition occurs due to the presence of the accumulator (integrator). Theoretically, it has infinite DC gain, so for a constant DC input, its internal state is unbounded. This presents a problem, since one cannot use an infinitely wide bus. The solution is to introduce some form of node saturation that clips the maximum or minimum signal level if overflow or underflow occur respectively. Clipping logic is located at the loop filter output in Figure 5.8, which also has an additional output indicating whether the filter has entered saturation. This feature is useful to determine whether the main synthesizer loop has locked, since an unlocked loop will eventually cause the integrator to overflow, thus saturating the loop filter. It is important to prevent any signal saturation during steady-state operation, because clipping

is a non-linear process although usually not as serious as limit cycles due to overflow. The dynamic range of the steady-state signal during normal operation (which is application specific) is used to set the appropriate bus widths. Ironically, during acquisition, intentional clipping is a benefit since it prevents the initial synthesizer phase error from growing too large. This preventative measure leads to faster acquisition without any performance penalty compared to a conventional analog PLL.

A large dynamic range in the filter avoids any loss of precision due to overflow and underflow. The consequence of maintaining a large dynamic range is the final filter output will also be large. In this application, the loop filter output drives a digital-to-analog (D/A) converter which has a limited resolution (i.e. 12-bit input word size). Reducing the output dynamic range can be done in a number of ways. One can simply truncate enough least significant bits (LSB's) so the remaining word width matches the D/A converter input size, but this produces quantization noise. The information lost in the LSB's can be recovered by remodulating these bits into a single bit using a technique described in Section 3.4.2.

Figure 5.9: Reducing the D/A dynamic range requirement through $\Delta\Sigma$ remodulation.

This is shown in Figure 5.9, where the $\Delta\Sigma$ modulator converts the truncated LSB's into a low resolution bitstream that is summed with the remaining MSB's. The quantization noise is spectrally shaped by the modulator so most of the noise power lies out of band. One may question that the use of $\Delta\Sigma$ modulation reintroduces quantization noise where earlier efforts were made to reduce it. However, the important point is the total noise power due to remodulating a few LSB's is much less than remodulating the entire word. For example, in this realization, the 18-bit loop filter output is converted to a 10-bit signal

by remodulating 8 LSB's. The amount of additional quantization noise introduced is dependent on the quantization level Δ used, which in this case is $2^8 = 256$. This corresponds to a noise power or variance equal to

$$E_{RMS}^2 = \frac{\Delta^2}{12} = \frac{(2^8)^2}{12} \quad (5.1)$$

which is $(2^{18-8})^2 = (1024)^2$ times less than if the entire 18-bit word were to be remodulated. The small amount of shaped quantization noise in Equation (5.1) has a negligible effect on the synthesizer phase noise, so additional filtering isn't necessary.

One final requirement is to convert the signed output of the loop filter into an unsigned value compatible with the 12-bit D/A converter (Analog Devices AD568). This amounts to adding an offset to the filter output to shift all negative values to corresponding positive values, as shown in Figure 5.9.

5.2.2 Digital-to-Analog Converter

Aside from the obvious need to convert the digital output of the DSP into a zero-order held analog signal, the D/A converter must also pass the modulation signal without significant distortion. This amounts to a number of necessary converter characteristics to guarantee correct operation of the modulator.

Since the D/A converter lies in the forward path of the main synthesizer loop, some non-linearities are able to be corrected through feedback. However, since the modulation data rate is greater than the loop BW, the corrective action of the loop is too slow so signal distortion is imminent. This implies that the D/A converter linearity must be high enough over the modulation bandwidth of interest. Another potential problem area is the introduction of glitches at the converter output due to input data skew or internal current source mismatch. This form of distortion will cause reference feed-through, since the converter changes state at the clock frequency of the DSP, which is identical to the reference frequency. Careful attention to the timing of the bits at the D/A converter input will minimize any external data skewing, while the design of the internal current source

switches must be symmetrical to reduce internal data skew.

5.2.3 Analog Integrator

Continuous-time integration of the D/A converter output serves a number of purposes. The integrator supplies the missing integration to make the synthesizer a second-order phase-locked loop (recall the $\Delta\Sigma$ FD output is a frequency error). This integration could have been realized in the digital domain, but there would still be a need for some filtering of the quantization noise at the D/A output. By replacing the output filter with a continuous-time integrator, one less filter is required. It also provides some open-loop gain adjustment to compensate for the unknown VCO sensitivity. Without any gain adjustment, the transmit equalizer would incorrectly compensate for the closed-loop modulation response as described in Section 3.4.4.2. An automatic tuning technique of the open-loop gain, using existing DSP, is described in [Bax98a], but implementation of the technique remains a topic for future research.

There are a number of architectures that can be used to realize a continuous-time integrator. In Chapter 4, the integration in the forward path of the $\Delta\Sigma$ FD was realized using a charge pump, which provides a simple interface to the differential PFD output. Acceptable matching of the current sources was possible by fully integrating the design onto a single silicon substrate. However, the D/A output in the GMSK modulator is a continuous zero-order held signal (i.e. it produces voltage steps, not pulses), so an integrator with active gain is used instead, as shown in Figure 5.10. Gain adjustment is provided through the variable resistor at the inverting input. An important consideration is the effect of the non-ideal op-amp on the performance of the integrator. In particular, one needs to define the required op-amp performance (open-loop gain, BW etc.) based on the GMSK modulator operating parameters. Most commercial op-amps use a single-pole compensation scheme, so the open-loop gain rolls off at a constant -20dB/dec after the pole frequency. To evaluate the effect of op-amp limitations on the integrator performance,

$$v_o(t) = \frac{-1}{RC} \int v_i(t) dt$$

Figure 5.10: Continuous-time integrator with variable gain and negative output clamp.

the op-amp is modeled as a single-pole lowpass filter with DC gain A_o and a unity-gain bandwidth (UGBW) ω_t . The frequency dependent op-amp open-loop gain is

$$A(s) = \frac{A_o}{1 + \frac{s}{\omega_b}} \quad (5.2)$$

where ω_b is the compensating pole frequency. An accurate expression for the response of the integrator in Figure 5.10 can be found by substituting for the non-ideal op-amp response, rather than assuming an infinite gain with unlimited BW. This leads to the integrator transfer function

$$H_{int}(s) = \frac{-1}{sRC} \cdot \frac{1}{1 + \left(1 + \frac{1}{sRC}\right)/A(s)} \quad (5.3)$$

which is the ideal integrator response $-1/sRC$ multiplied by the effect of the non-ideal op-amp. Now substituting for $A(s)$ in Equation (5.3) gives

$$H_{int}(s) = \frac{-1}{sRC} \cdot \frac{1}{1 + \frac{1}{A_o} \left(1 + \frac{1}{sRC}\right) \cdot \left(1 + \frac{s}{\omega_b}\right)} \quad (5.4)$$

Unity gain occurs at $\omega_t = A_o \omega_b$ so Equation (5.4) can be written as

$$\begin{aligned}
H_{int}(s) &= \frac{-1}{sRC} \cdot \frac{1}{1 + \frac{1}{A_o} + \frac{1}{sRCA_o} + \frac{1}{\omega_t RC} + \frac{s}{\omega_t}} \\
&= \frac{-1}{RC} \cdot \frac{1}{\left(\frac{1}{\omega_t}\right)s^2 + \left(1 + \frac{1}{A_o} + \frac{1}{\omega_t RC}\right)s + \frac{1}{A_o RC}}
\end{aligned} \tag{5.5}$$

Equation (5.5) shows that the op-amp limitations transform the ideal integrator to a second-order response. If the open-loop gain A_o is too small, the integrator will suffer from a finite DC gain. If there is insufficient BW, the second closed-loop pole will shift the response from -20dB/dec to -40dB/dec within the desired BW, increasing the gain error near the UGBW.

An important point to consider is that the required op-amp BW is set by the modulation BW, not the synthesizer reference frequency. This implies that accurate integration of the D/A converter output is only necessary for the modulation signal, while outside that BW, the response can naturally attenuate due to the op-amp finite UGBW. The limited BW of the op-amp provides additional filtering to reduce the system phase noise and reference feed-through. The only restriction is that the op-amp BW must be much larger than the synthesizer closed-loop BW or the effect of the extra pole (see Equation (5.5)) must be considered in the stability analysis.

For GSM, a suitable modulation BW is 1 MHz, which ensures that the GMSK signal is not distorted due to the analog integrator. If the peak gain error is to remain less than 1 dB, the op-amp must have an open-loop gain $A_o=87$ dB and a UGBW=2 MHz as shown in Figure 5.11.

During loop acquisition, a potential problem exists that the linearized models do not predict. When power is applied, the initial state of the integrator may be close to the negative power rail, since it has a bipolar output. The models predict that the VCO would oscillate at its minimum frequency, after which the loop would eventually lock. Practically, many VCOs cease to oscillate if the tuning input is below some minimum voltage. Below this critical voltage, the VCO tuning varactor becomes forward biased, and

Figure 5.11: Continuous-time integrator response with non-ideal op-amp ($A_o=87$ dB, $UGBW=2$ MHz).

the VCO oscillation ceases. Without a valid VCO output, the discriminator output is unpredictable, as it will produce either a stream of ones or zeros. If it produces a stream of ones (representing positive saturation), the synthesizer will erroneously produce an error signal to reduce the VCO frequency and the loop will never lock. Restricting the minimum integrator output voltage to a level compatible with the VCO input will mitigate this problem. This will not interfere with any modulation data since the channel frequency would maintain a constant DC offset at the VCO input. Placing a diode in the feedback path of the op-amp integrator (see Figure 5.10) effectively clamps the output voltage whenever it drops too low, since the diode will be forward biased. The diode is reverse biased during normal operation, which simply adds a small depletion capacitance (i.e. $C_j \ll C_{int}$) in parallel with the integrating capacitor without any adverse effect. If the integrating capacitor C_{int} in Figure 5.10 is too small (RC time constant determined by integrator gain K_i), the effect of the diode depletion capacitance can be significant. Since

C_j varies with reverse bias (i.e. integrator output voltage), the effective integrator gain K_i will vary with channel frequency, as the VCO tuning voltage changes. This will adversely make the synthesizer open-loop gain K dependent on the channel frequency, which makes the loop dynamics channel dependent.

5.2.4 Voltage-Controlled Oscillator

A discrete VCO module with a centre frequency in the 2GHz range is used in the modulator prototype. Integrated VCO's with on-chip LC resonators [Daup97] have been investigated, but their development has not matured enough to warrant their use in this prototype modulator. Current research in low-power transceiver design indicates that integrated VCO's will ultimately offer significant power and cost savings in the future.

Several issues must be considered before choosing a commercial VCO to use in the GMSK modulator. The foremost parameters are its phase noise and input sensitivity, which both affect the noise at the transmitter output. Analysis of any synthesizer will show that the phase noise outside the loop BW is dominated by the VCO open-loop phase noise. In this region, there is no noise suppression, so this phase noise must be low enough to meet the desired specifications. Inside the loop BW, any noise is suppressed by an amount depending on the synthesizer order, closed-loop BW etc. The GSM standard specifies a transmit phase noise less than -105dBc/Hz at a 400KHz offset from the carrier [GSM98], which implies that the VCO must perform even better, to allow for all other wideband noise sources in the loop. External noise at the tuning input can be amplified by a significant amount if the VCO sensitivity K_v is high. Most of this voltage noise originates in the active loop filter and gets converted to an equivalent phase noise by the VCO. In this modulator architecture, a VCO sensitivity of around 20MHz/V is used to balance the noise amplification and dynamic range requirement of the loop filter (i.e. for a fixed modulation frequency deviation, loop filter dynamic range is inversely proportional to VCO sensitivity). With these considerations in mind, a commercial VCO whose phase noise is shown in Figure 5.12 was chosen. The measured operating frequency of this VCO

Figure 5.12: Open-loop phase noise of Z-COMM model V613ME04 VCO.

was 1.7GHz to 2GHz which is suitable for the GSM DCS-1800 and DCS-1900 standards.

Complete noise analysis is reported in Section 5.4, which combines the noise from all sources to determine their effect on the output phase noise of the synthesizer.

5.3 Modulation Data Path

The modulation path of the transmitter in Figure 5.1 is composed of the blocks shown in Figure 5.13. The 1-bit data symbols, representing a ± 1 , are convolved with the transmit filter impulse response to produce an equalized GMSK filtered signal, sampled at f_r . The resulting multi-bit word is then summed with a constant channel frequency offset (200KHz spacing) and the output is $\Delta\Sigma$ modulated to reduce the resolution prior to

controlling the modulus input of the $\Delta\Sigma$ FD. The strategies used to realize the modulation path are described in the following sections.

Figure 5.13: Digital modulation data path.

5.3.1 Digital Transmit Filter

There are two options available when realizing the transmit filter that arise from the modulation path being outside the main synthesizer loop. The entire transmit filter (including the $\Delta\Sigma$ modulator) may be realized in real-time software that produces the final output at the oversampled rate. This is the easiest and most flexible solution and can be used to evaluate various transmit filter architectures and the non-ideal effects of finite amplitude resolution. What is lost in this approach is any sense of area or power that the filter may consume when realized in hardware in a given integrated circuit technology. Alternatively, the filter may be realized for hardware exploration in an FPGA or ASIC. While a complete hardware solution would be able to provide more useful data, a software solution was chosen to give filter topology flexibility without investing significant design time for implementation.

Combining the Gaussian and equalization filter responses produces a transfer function of the general form

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2} + \dots + b_n z^{-n}}{1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_n z^{-n}} \quad (5.6)$$

Equation (5.6), when realized in direct form using fixed-point arithmetic, results in an IIR structure, which complicates the hardware as described in Section 5.2.1. If the filter has a finite impulse response (or can be truncated and windowed to resemble one), a simpler ROM based architecture, similar to that used for the quantization noise filter in Section 5.2.1, can be used. This approach is desirable since it reduces any real-time arithmetic, performed at the oversampled rate f_r , to a much simpler table look-up of partial sums (refer to Figure 5.4). The drawback of such a filter is that only zeros can be realized (i.e. a_n terms are zero), so a much higher order is required to approximate a given filter response. With future hardware realization in mind, a software version of a ROM based FIR transmit filter was used to Gaussian filter the symbols while equalizing the effect of the limited synthesizer closed-loop BW.

Adopting the GSM standard, results in significant ISI from the adjacent symbols on either side of the current one due to the narrow Gaussian filter bandwidth ($BT = 0.3$). This effect is clearly seen in Figure 5.14 where the Gaussian filtered symbols are smeared

Figure 5.14: GSM Inter-symbol interference between (a) individual symbols and (b) the combined effect on a symbol sequence.

in time across adjacent symbols. This implies that the desired frequency trajectory depends on the sequence of three symbols, not just the current one, which must be taken into account when indexing the ROM filter. The amplitude resolution of the look-up table values was set to 8 bits to ensure the desired frequency trajectory does not deviate too far from the ideal GMSK filtered data.

Using the GSM modulator parameters defined in Section 5.1, and input data symbols of ± 1 , results in a dynamic range increase from 1 bit to 16 bits at the transmit filter output in Figure 5.13. This multi-bit signal must be added to the 7-bit channel offset (i.e. 200KHz GSM channel spacing) to form the complete modulated (and scaled) carrier signal. This high resolution signal is remodulated into a lower resolution signal prior to injecting it into the $\Delta\Sigma$ frequency discriminator, using a digital $\Delta\Sigma$ modulator as shown in Figure 5.15 and discussed in the next section.

5.3.2 Digital MASH $\Delta\Sigma$ Modulator

As in the case of the transmit filter, the digital $\Delta\Sigma$ modulator in Figure 5.15 can be realized entirely in software or implemented in hardware. If a hardware approach is chosen, the recursive nature of the modulator forbids the use of any ROM based architecture. The

Figure 5.15: Digital second-order MASH $\Delta\Sigma$ modulator block diagram.

same limitation applied when trying to realize the transmit filter in its original IIR form. However, the use of a cascaded modulator architecture (i.e. MASH structure), which was determined earlier as the best choice for extended dynamic range, is ideal for pipelining, since there isn't any global feedback. Pipelining the modulator architecture relaxes the speed requirements of the adders at the expense of some latency (i.e. processing delay), since each output may now take a number of clock cycles to be computed. The reduced adder speed requirement results in lower overall power consumption [Lu93], [Jou97].

A software (SIMULINK[®]) model of the block diagram in Figure 5.15, which included the effects of finite resolution (i.e. finite internal bus width) was used to realize the second-order MASH $\Delta\Sigma$ modulator. While this approach proved adequate in a simulation environment, it presented a problem in actual hardware testing. The reason is due to the finite $\Delta\Sigma$ modulator sequence length that is produced during a simulation. If this same sequence were to be injected in real time into the GMSK modulator repeatedly, a discontinuity would exist during the transition from one sequence to the next. The phenomenon behind this behaviour is due to the disruption in the internal states of the digital $\Delta\Sigma$ modulator, which by definition should be chaotic. That is, appending a $\Delta\Sigma$ output sequence to itself such that the first and last values coincide, does not imply that the internal states match. For example, if a $\Delta\Sigma$ modulator with a 1-bit comparator produces the sequence shown in Figure 5.16 for a given input, its initial and final internal states are different (i.e. $S_1 \neq S_8$). Splicing two of these sequences together (by overlapping the last and first output samples) introduces a state change which would not occur if the original sequence were left to continue. Although finite sequence lengths will affect the performance of the GMSK modulator prototype, they can still be used for hardware testing. The solution is to use a long enough sequence length that ensures the GMSK modulator has reached steady state, and then trigger the test instrument to sample the output. This does not remove the finite sequence effect, but rather allows one to observe

Figure 5.16: Discontinuities introduced by splicing finite length $\Delta\Sigma$ modulator outputs.

the system during steady-state periods.

5.4 Noise Analysis

This section examines the influence of various noise sources within the synthesizer and their effect on the phase noise at the output. There are two operating modes to consider — local oscillator mode without modulation, and transmit mode with GMSK modulation applied. Conventional noise analysis is used to observe the effect of various synthesizer noise sources (including $\Delta\Sigma$ frequency discriminator noise) on the unmodulated carrier signal. This analysis provides a measure of the total output phase noise which can be compared to the limits imposed by the GSM standard. Once an acceptable phase noise spectral density is achieved, the next step is to investigate the effect on the phase noise, while modulating the synthesizer with GMSK filtered data. The key point to consider is whether adding modulation adversely affects the phase noise within the modulation bandwidth or elsewhere. Inside the modulation BW, noise will directly affect the achievable signal-to-noise ratio, so the total noise power must be constrained to some level determined by the standard. Outside the modulation BW, the issue is whether the noise

density (which was acceptable without modulation) will be increased due to the modulation and potentially violate the maximum tolerable noise level.

Using the synthesizer block diagram model of Figure 3.12 (redrawn with the dominant

Figure 5.17: Noise model for GMSK modulator using linear $\Delta\Sigma$ FD model.

noise sources S_{xx} visible in Figure 5.17), the output-referred transfer function of each noise source can be computed. Calculation of the actual output noise amounts to defining the spectral density of each noise source and referring them to the synthesizer output. Assuming that each noise source is uncorrelated with any other, the total output noise is simply the sum of all the independent output-referred noise sources. Figure 5.18 shows the spectral density of each noise source and its corresponding transfer function to the output. The $\Delta\Sigma$ modulator and $\Delta\Sigma$ FD noise spectra both exhibit the expected highpass shape that is periodic about integer multiples of the sampling rate f_r . They are both effectively

Figure 5.18: Equivalent block diagram of synthesizer noise sources.

lowpass filtered, attenuating the quantization noise. The VCO phase noise is highpass filtered, which suppresses the noise within the loop BW, but beyond that the noise is unaltered. The final noise source, which originates at the analog integrator output, is bandpass filtered by the synthesizer loop. The integrator noise combines the op-amp voltage and current noise densities and the passive component noise (e.g. resistors). When this noise is referred to the output of the synthesizer, it assumes a lowpass shape with a uniform spectral density within the synthesizer loop BW.

It becomes clear that the choice of synthesizer order, closed-loop cut-off frequency f_o and reference frequency f_r affect the noise transfer function and ultimately control the output phase noise. These relationships are qualitatively listed in Table 5.4, which indicate that a higher order of PLL and higher reference frequency reduce the total phase noise.

The actual choice of these parameters is governed by the modulation requirements stated

Table 5.4: Influence of synthesizer parameters on individual noise sources.

Noise source	PLL order $\uparrow$	cut-off $f_o \downarrow$	reference $f_r \uparrow$
$S_{\Delta\Sigma}(f)$	reduced	reduced	reduced
$S_{VCO}(f)$	reduced	increased	no effect
$S_{\Delta\Sigma FD}(f)$	reduced	reduced	reduced
$S_{INT}(f)$	reduced	increased	no effect

earlier in Section 5.1. The only conflicting parameter is the closed-loop cut-off frequency f_o , which is related to the open-loop BW. Both $\Delta\Sigma$ noise sources will benefit from a reduction of f_o , since most of the noise is shaped and exists at higher frequencies. However, the VCO and analog integrator noise will suffer from inadequate close-in noise suppression with small values of f_o , so a compromise must be made.

Assuming that each of the noise sources in Figure 5.17 are independent of each other, the overall noise spectral density in [rad^2/Hz] can be expressed as

$$S_{\phi}(f) = S_{\phi, \Delta\Sigma}(f) + S_{\phi, VCO}(f) + S_{\phi, \Delta\Sigma FD}(f) + S_{\phi, INT}(f) \quad [\text{rad}^2/\text{Hz}] \quad (5.7)$$

where each of the output referred noise sources are defined by

$$\begin{aligned}
 S_{\phi, \Delta\Sigma}(f) &= S_{\Delta\Sigma}(f) \left| \frac{-H_{DSP}(f) \cdot H_{DAC}(f) \cdot H_{INT}(f) \cdot H_{VCO}(f)}{1 + H_{OPEN}(f)} \right|^2 \\
 S_{\phi, VCO}(f) &= S_{VCO}(f) \left| \frac{1}{1 + H_{OPEN}(f)} \right|^2 \\
 S_{\phi, \Delta\Sigma FD}(f) &= S_{\Delta\Sigma FD}(f) \left| \frac{H_{DSP}(f) \cdot H_{DAC}(f) \cdot H_{INT}(f) \cdot H_{VCO}(f)}{1 + H_{OPEN}(f)} \right|^2 \\
 S_{\phi, INT}(f) &= S_{INT}(f) \left| \frac{H_{VCO}(f)}{1 + H_{OPEN}(f)} \right|^2
 \end{aligned} \quad (5.8)$$

The spectral density for each noise source was obtained from manufacturers data (e.g. VCO, op-amp etc.) and from simulation (i.e. Hspice[®] linear AC noise analysis). Plots of these spectra using the synthesizer parameters specified in Section 5.1 are shown in Figure 5.19. From this plot, it is clear that within the 30KHz loop BW, the analog integrator (i.e.

op-amp) noise dominates. For intermediate frequencies, the $\Delta\Sigma$ quantization noise from the $\Delta\Sigma$ FD and digital modulator is prominent while high frequency noise is dominated by the VCO. The total synthesizer phase noise is clearly compliant with the GSM transmit

Figure 5.19: Simulated phase noise of GSMK modulator.

phase noise mask, as shown on the Figure.

Earlier noise analysis produced a measure of the expected phase noise (Figure 5.19) with the synthesizer loop locked but without applying modulation (i.e. a local oscillator). In transmit mode, the carrier is modulated with the GMSK filtered data. Whether this modulation process affects the synthesizer phase noise can be determined by decomposing

Figure 5.20: Time-domain modulator output signal composed of equivalent GMSK phase modulation and synthesizer phase noise.

the RF output signal into modulation and noise components as in Figure 5.20. Thus the synthesizer output can be expressed as

$$out(t) = \cos(2\pi f_c t + \phi_{mod}(t) + \phi_n(t)) \quad (5.9)$$

where $\phi_{mod}(t)$ is the equivalent phase of the GMSK modulation (recall GMSK is a frequency modulation). Equation (5.9) can be expanded using a trigonometric identity as

$$out(t) = \cos(2\pi f_c t + \phi_{mod}(t)) \cdot \cos(\phi_n(t)) - \sin(2\pi f_c t + \phi_{mod}(t)) \cdot \sin(\phi_n(t)) \quad (5.10)$$

Assuming that the synthesizer phase noise is small (i.e. $|\phi_n(t)| \ll 1$), Equation (5.10) reduces to

$$\begin{aligned} out(t) &\approx \cos(2\pi f_c t + \phi_{mod}(t)) - \sin(2\pi f_c t + \phi_{mod}(t)) \cdot \phi_n(t) \\ &\approx \cos(2\pi f_c t + \phi_{mod}(t)) + \cos(2\pi f_c t + \phi_{mod}(t) + 3\pi/2) \cdot \phi_n(t) \end{aligned} \quad (5.11)$$

To obtain the power spectral density, it is convenient to take the Fourier transform of the autocorrelation of Equation (5.11), since they are transform pairs (i.e. $\mathcal{F}\{R(t)\} \leftrightarrow S(f)$). The synthesizer phase noise is uncorrelated with the modulation signal, so the cross-correlations between them is zero. This also implies that the autocorrelation of their product (second term in Equation (5.11)) is the product of their independent autocorrelations [Pete95]. Thus the autocorrelation of Equation (5.11) is

$$R_{out}(t) = R_{mod}(t) + R_{mod}(t) \cdot R_{\phi_n}(t) \quad (5.12)$$

Finally, the power spectral density of the modulator output is simply the Fourier transform

of Equation (5.12) which is

$$\begin{aligned} S_{out}(f) &= \mathcal{F}\{R_{out}(t)\} \\ &= S_{mod}(f) + S_{mod}(f) * S_{\phi_n}(f) \end{aligned} \quad (5.13)$$

where '*' denotes the convolution operator. From Equation (5.13), it is clear that modulating the synthesizer produces the desired modulation spectrum $S_{out_m}(f)$ in addition to converting the original phase noise $S_{\phi_n}(f)$ into $S_{out_m}(f) * S_{\phi_n}(f)$. The convolution of the modulation with the phase noise either raises the phase noise level or converts any spurious tones into an equivalent phase noise as shown in Figure 5.21. From

Figure 5.21: Output spectrum of (a) unmodulated carrier and (b) GMSK modulated carrier.

a spurious viewpoint, directly modulating a synthesizer is beneficial since all discrete spurs are convolved into phase noise. As long as the total phase noise is below the required limits, applying modulation has no adverse effects on the synthesizer noise performance.

It is interesting to compare the direct modulation approach used here to conventional up-conversion using mixers [Stet95]. In the mixer case, the local oscillator (LO) spurs are

convolved into phase noise by the baseband modulation signal. However, the mixer doesn't have infinite isolation from its LO port to its output, so some of the LO noise spurs are fed through to the output. Careful choice of the synthesizer reference and IF frequencies, in conjunction with the output filter BW, must be made to ensure the spurious noise is kept within the desired limits.

5.5 Modulator Performance

Evaluating the performance of a transmitter consists of two main criteria — qualifying how faithfully the modulation is subsequently reproduced at the receiver, and its noise performance in the absence of a modulating signal. The received modulation signal is compared in Section 5.5.1 to the simulated results, which represent the best possible performance achievable with this architecture. In this case, the system model developed in Chapter 3 was used to include the effects of $\Delta\Sigma$ quantization noise and finite resolution in the DSP. Noise is discussed separately in Section 5.5.2, although it will influence the transmitter's ability to modulate the carrier. Evaluating the transmitter, while disabling the modulation, allows the true noise floor to be determined. Then the key point to consider is whether there is adequate signal-to-noise ratio (SNR) during modulation to accurately reproduce the data being transmitted.

5.5.1 GMSK Transmitter Performance

In this section, the performance of the transmitter under optimal and non-ideal conditions is evaluated. The results are based on the parameter set developed for the GSM modulator example in Section 5.1. Before examining the modulator itself, it is useful to establish a baseline performance extracted from ideal GMSK modulation of a carrier in a noise free system. Simulation results for such an ideal modulator with GSM modulation are shown in Figure 5.22(a). The GSM transmit spectral mask is superimposed on the spectrum to clearly show that the RF signal power is contained within the desired levels of the GSM standard. Optimal simulated results (i.e. the equalization exactly compensates for the PLL

closed-loop attenuation) for the proposed GMSK modulator are shown in Figure 5.22(b). Comparing the ideal RF output spectrum to the optimal GMSK modulator spectrum,

Figure 5.22: Output power spectral density for an (a) ideal GSM modulated carrier and (b) simulated GMSK modulator with optimal parameter set.

shows that the residual noise from the $\Delta\Sigma$ quantization in the digital modulator and the $\Delta\Sigma$ FD raises the noise floor at offsets greater than 400KHz from the carrier. Even though the GMSK modulator introduces additional noise, the resulting RF spectral power is still within the required limits set by the GSM standard.

Figure 5.23 shows the measured output power spectrum for a GSM modulated carrier at 1.8655GHz obtained with an HP8593E spectrum analyzer. The GSM transmit spectral mask is included to verify the modulated carrier power is contained within the required limits. The measured noise level, at frequency offsets greater than 400KHz, is higher than the simulated spectrum in Figure 5.22(b), due to the omission of any circuit noise sources other than $\Delta\Sigma$ quantization noise in the simulated results. However, the measured noise at offsets greater than 400KHz is still 10dB below the required GSM mask so this doesn't compromise the modulator performance.

Modulation accuracy of the transmitter is determined at the receiver by demodulating the transmitted carrier and extracting the in-phase and quadrature signals (I & Q) as illustrated in the test set up of Figure 5.24. The received I and Q signals, along with an

Figure 5.23: Measured output power spectrum of GMSK modulator with a 1.8655GHz carrier frequency.

ideal reference signal, can be mapped into a constellation by plotting I versus Q as in Figure 5.25. This format is simply a plot of the modulation vector (i.e. amplitude and phase) sampled at the decision making times. From this plot, a number of observations can be made. First, the GSM signal is by definition a constant amplitude modulation scheme,

Figure 5.24: Vector modulation analyzer test set up.

so the plot of I versus Q transcribes a circle. Second, GSM modulation uses minimum frequency shift keying which produces orthogonal I and Q signals. This simply means that there are four distinct regions on the I-Q constellation separated by $\pi/2$ radians of phase per symbol. Note that the ideal GSM signal (reference plot in Figure 5.25) doesn't produce four distinct constellation points, but rather they are spread in phase due to the ISI introduced by the narrow Gaussian filter bandwidth ($BT = 0.3$). The measured constellation in Figure 5.25 spreads even further than the reference due to phase error in the transmitter. This phase error results from non-ideal carrier frequency trajectories during modulation. Spreading of the constellation causes the symbols to be closer together and results in a higher probability of error during symbol detection. If the spreading is

Figure 5.25: Reference and measured constellations of GSM modulated carrier ($BT=0.3$).

very severe, the receiver may lose symbol lock, where it cannot demodulate the signal or recover any information.

In the GMSK modulator architecture, the open-loop gain K is the only unknown loop parameter that can cause potential problems if it differs from the value used to synthesize the equalization filter. As discussed in Chapter 3, any mismatch of the open-loop gain from the expected value, will result in an under or over compensated modulation transfer function (see Figure 3.26). This will not have a significant effect on the RF spectrum,

since the modulation BW is much larger than the closed-loop BW, but the spectrum will be slightly distorted depending on the amount of gain error. Thus an open-loop gain error introduces a DC gain error over most of the modulation BW. This gain error directly affects the frequency deviation, as well as introducing additional ISI beyond that from the GSM Gaussian filter.

Received I and Q eye diagrams were measured using the vector modulation analyzer test set-up in Figure 5.24 for three cases — optimally matched with 0% gain error and with $\pm 20\%$ open-loop gain error. These results were compared to the simulated results to

Figure 5.26: Simulated (a) and measured (b) I and Q eye diagrams with 0% open-loop gain error.

confirm the predicted modulator behaviour under these conditions. Figure 5.26(b) illustrates the optimally matched case where the hardware open-loop gain has been adjusted to compensate for the unknown VCO sensitivity K_v . Comparing this to the simulated plot in Figure 5.26(a), it is clear that the eyes have well defined openings with compact crossover points. These characteristics reduce the error during symbol detection leading to better system performance. Figure 5.27 depicts the case for a deliberate +20% gain error. As expected, the modulator performance is compromised, and this is evident in the eye diagrams which suffer from a slightly closed eye with ill defined crossover points. This makes the receiver more sensitive to timing errors since the eye remains open for less

Figure 5.27: Simulated (a) and measured (b) I and Q eye diagrams with +20% open-loop gain error.

Figure 5.28: Simulated (a) and measured (b) I and Q eye diagrams with -20% open-loop gain error.

time. The final case is shown in Figure 5.28, where the gain error has now been reduced to -20%. The results are similar to those with a positive gain error, but the quality of the eye has been further compromised. The reduction of the eye opening reduces the available noise margin during symbol detection in addition to the timing sensitivity from spreading of the zero crossings.

Results from experimental measurements of the modulator in transmit mode are comparable to the simulated results, although the hardware prototype suffered from a lower signal-to-noise ratio than expected. This was evident in the quality of the measured eye diagrams and constellation plots of the modulator with no mismatch. Since the signal levels (i.e. frequency deviation for GMSK) are well defined in the digital transmit filter, the SNR degradation is due to a higher noise floor than expected as explained in the next section.

5.5.2 Synthesizer Performance

Synthesizer performance pertains to the GMSK transmitter operating in its unmodulated state such that only a carrier is present. In this mode, the transmitter noise performance can be evaluated to ensure it is low enough to maintain an adequate SNR in-band during modulation and prevent any interference out-of-band. This mode is also active when the channel frequency is changed (e.g. either user initiated or during frequency hopping), since transmission of data is disabled.

Figure 5.29 shows an approximate measure of the closed-loop noise performance made by observing the output spectrum of the synthesizer with an HP8593E spectrum analyzer. From this figure, the suppression of the close-in phase noise up to the 30KHz closed-loop BW is readily visible. The actual phase noise can be estimated, assuming the small angle criteria is not violated, by measuring the relative noise level with respect to the carrier and compensating for the finite resolution bandwidth (RBW) of the spectrum analyzer. In this case, the close-in noise level is at -40dBc and the RBW is 1 KHz so the

estimated phase noise is

$$\begin{aligned}\mathcal{L}(f) &= -40\text{dBc} - 10\log(1\text{KHz}) \\ &= -70\text{dBc/Hz}\end{aligned}$$

Figure 5.29: Measured synthesizer output spectrum.

The out-of-band noise response is shown in Figure 5.30. An important parameter in any synthesizer is the amount of reference suppression, and in this architecture it is even more important since the VCO is controlled by the DSP through a digital-to-analog converter. If any of the digital switching noise couples to the VCO tuning port, reference spurs will exist in the RF output. From Figure 5.30, the measured reference feed-through was less than -76dBc. Due to the limited dynamic range of the spectrum analyzer, spurs less than -80dBc are not discernible. The only accurate way to measure these spurs (and close-in phase noise) is by performing a true phase noise measurement which is presented next.

An accurate measure of the synthesizer phase noise was accomplished with an HP3048A phase noise measurement system. This system can achieve a much higher

Figure 5.30: Measured synthesizer spurious noise.

dynamic range, and thereby a lower noise floor, by mixing down the RF carrier to DC. This is achieved by incorporating the synthesizer within a PLL and measuring the resultant noise from the phase detector output. This technique also allows the close-in phase noise to be measured where earlier, with a spectrum analyzer, this was not possible.

Initial phase noise measurements revealed that the hardware prototype had a higher in-band noise floor than that expected from simulation (see Figure 5.19). This was understandable, since the measured results of the hardware discriminator in Chapter 4 exhibited less than ideal noise shaping. Bearing this in mind, the expected synthesizer phase noise in Figure 5.19 was re-computed using the *measured* $\Delta\Sigma$ FD noise (other noise sources remain simulated) to produce the simulated synthesizer phase noise in Figure 5.31. This closely matches the measured phase noise in Figure 5.32 within a few dB, which validates the PLL loop models used in the noise analysis. Note that some open-loop gain error due to manual adjustment of the open-loop gain causes some second-order

Figure 5.31: Simulated phase noise of GMSK modulator using *measured* $\Delta\Sigma$ FD noise.

peaking in the measured phase noise.

Inspection of the measured phase noise in Figure 5.32 shows an in-band noise level of -74dBc/Hz and a reference spurious suppression of -80dBc , which is close to the estimated values obtained from the spectrum analyzer. In both the simulated and measured phase noise plots, the GSM spectral mask is included to verify that the noise performance (even with higher in-band phase noise) is within the desired limits. The consequence of increased close-in phase noise is a reduction of the signal-to-noise ratio while in transmit mode. That is, the signal power, which is set by the modulation standard, is constant so the increased noise degrades the SNR. As long as the modulator SNR is high enough to reliably allow detection of the symbols at the receiver, the BER will be acceptable.

Figure 5.32: Measured synthesizer phase noise.

The synthesizer transient response, characterized by its switching speed, was measured by allowing it to reach steady state and then injecting an input frequency step large enough to force it to lose phase lock. The settling time was measured using an HP5372A frequency and time interval analyzer that was triggered at the same time as the input step. For this test case, the initial carrier frequency was 1.86GHz and an input frequency step of 6MHz was used. Since the input frequency step is much less than the carrier frequency, the required resolution for the frequency and time interval analyzer is very high. This problem was circumvented by repeating the input step and averaging the results, which are illustrated in Figure 5.34. Comparing the measured switching speed of 270 μ s to the 250 μ s simulated result in Figure 5.33 shows a good degree of correlation, although some error is evident due to the limited resolution of the measurement.

An interesting characteristic of this modulator is that it doesn't exhibit the classic second-order transient response seen in the literature. The reason for this is that during acquisition (in this case caused by the synthesizer losing lock after the input step), the phase error is large enough to exercise the entire dynamic range of the DSP. Once this

limit is reached (either positive or negative), the DSP saturates to a fixed value until the phase error reduces through feedback in the loop. However, from the modulator block diagram in Figure 5.1, it is clear that the DSP output is converted to an analog signal and then integrated in the charge pump. This effectively enforces a maximum switching speed, similar to slew-rate limiting in an amplifier, since integrating a DC value produces a ramp. This non-linear behaviour dramatically alters the transient response as seen in the simulated and measured frequency versus time plots. Note that the switching speed can be

Figure 5.33: Simulated synthesizer switching speed for an input frequency step.

improved by increasing the dynamic range of the DSP (with a corresponding increase in DSP complexity) so a compromise must be made dependent on the desired application.

The 270 μ s measured switching speed was found to be similar to the simulated value. The effect of the DSP saturating is a slewing of the VCO tuning input. This limitation can

HP 5372A Frequency And Time Interval Analyzer

TVar: Frequency C 15 Jul 1999 13:27:10
 Δ Mkr x: 269.698395 us C evts N/A
 y: 6.16 MHz

Figure 5.34: Measured synthesizer switching speed for a 6MHz frequency step.

be improved by dynamically altering the loop parameters during channel switching and restoring them after phase lock is achieved to provide adequate noise control.

This chapter presented the design and implementation of a GMSK modulator that is suitable for wideband modulation. The flexible architecture can be used to realize many constant-envelope digital modulation schemes due to the digital transmit filter and digital loop parameters, which are both easily modified. To illustrate the feasibility of this architecture, a BiCMOS $\Delta\Sigma$ FD chip, DSP gate array and several discrete components were combined to form a complete GSM transmitter. Measured eye diagrams showed wideband modulation was possible through compensation of the synthesizer closed-loop BW.

Chapter 6

Conclusion

A new wideband transmitter architecture was presented in this thesis that is suitable for continuous-phase constant envelope modulation schemes. The technique uses direct modulation of a high resolution $\Delta\Sigma$ frequency discriminator based synthesizer to generate the RF signal without conventional up-conversion. This reduces the complexity and eliminates many of the analog problems found in conventional GMSK modulators. The resulting transmitter is primarily digital in nature and is more suitable for monolithic integration.

Digital equalization of the synthesizer closed-loop response makes wideband modulation possible (with respect to the PLL bandwidth), while simultaneously achieving good noise performance. This permits the use of a narrower loop bandwidth to attenuate noise, while equalization effectively widens the modulation bandwidth to accommodate higher data rates. Matching between the digital equalizer and synthesizer closed-loop response is not a significant issue in this architecture, since the synthesizer response (except for open-loop gain) is defined by digital parameters and is therefore predictable.

System level models of the complete transmitter were developed to verify the viability of the architecture for use as a general purpose GMSK modulator. Specific modulator applications only require modification of the transmit filter and synthesizer digital loop parameters to accommodate other modulation schemes. This feature alone makes this modulator architecture very flexible in that its performance can be completely redefined by simply exchanging one set of digital parameters for another. This can be done at the application level to accommodate different standards or dynamically to alter the

modulator performance in real time. An example of the latter case is to alter the PLL loop dynamics to improve acquisition and after it locks, restore the parameters for better noise performance.

Using the GSM standard (DCS-1800) as an example application, an experimental prototype modulator was designed and implemented using a fully integrated version of the $\Delta\Sigma$ frequency discriminator. The $\Delta\Sigma$ FD is a key part of the overall modulator, since it must perform analog-to-digital conversion of the high-frequency modulated RF signal. The discriminator was realized in a $0.8\mu\text{m}$ BiCMOS technology and consumed 75mW from a 3 V power supply. High speed operation is ensured by using a fully differential ECL/CML architecture, and the current distribution through various sections of the chip has been optimized for low power operation. Using these design techniques, the measured maximum RF input frequency was 2.5GHz, which is a more than adequate bandwidth for the modulation standards operating in the 2GHz band. The DSP in the modulator prototype was implemented in a programmable gate array for design flexibility, although it can easily be realized in any standard integrated circuit technology. This allowed different DSP architectures to be evaluated without having to go through the IC fabrication process. The remainder of the GMSK modulator was realized in discrete form on a printed circuit board.

Through experimental verification, it has been shown that the modulator architecture is feasible for use as a general purpose wideband GMSK transmitter. From a performance perspective, the hardware GMSK modulator test results revealed a higher in-band phase noise than anticipated from simulation results. This excess noise has a direct effect on the achievable SNR at the modulator output and results in a degradation of the received eye diagrams. The additional phase noise was due to less than ideal noise shaping in the integrated $\Delta\Sigma$ frequency discriminator. Since the discriminator lies in the feedback path of the modulator, any non-ideal effects cannot be corrected and appear at the output of the modulator. The lack of noise shaping in the discriminator output bitstream was traced to circuit noise originating from its internal phase-frequency detector. Lowering the noise floor of the reference signal f_r improved the performance somewhat, since it couples directly into the PFD and exhibits the same characteristics as internally generated noise.

6.1 Future Research

The results of this work verified the feasibility of a new wideband GMSK modulator architecture, but also open up opportunities for improved circuit performance and future research areas.

As mentioned previously, the modulator characteristics are defined in DSP and exploitation of this feature is an area of future research. For example, compensation for the unknown VCO sensitivity K_v could be achieved by using existing DSP to measure the open-loop gain, which is proportional to K_v , and digitally correcting for any deviation from the expected value. This effectively would allow any VCO to be used without initial tuning of the open-loop gain, or subsequent adjustment due to component drift over time or temperature.

Another area for future research deals with adaptive signal processing, where the loop dynamics can adapt to a certain operating mode. A classic example, applicable to this modulator architecture, is to alter the loop dynamics to decrease the acquisition time while switching channels. Once phase lock is achieved, the loop dynamics can be modified to obtain the best noise performance.

The suppression of discrete spurs arising from periodicity in the $\Delta\Sigma$ noise shaping in both the $\Delta\Sigma$ FD and the digital $\Delta\Sigma$ modulator in the transmit path is another area to investigate. Since these non-ideal effects are channel and operating mode dependent, the offending spur frequencies can be computed and an effort made to attenuate them through adaptive digital filtering.

References

- [Arda87] Sasan H. Ardalan and John J. Paulos, "An Analysis of Nonlinear Behaviour in Delta-Sigma Modulators", *IEEE Trans. on Circuits and Systems*, vol. CAS-34, no. 6, pp. 593-603, June 1987.
- [Bair94] Rex T. Baird and Terri S. Fiez, "Stability Analysis of High-Order Delta-Sigma Modulation for ADC's", *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 41, no. 1, pp. 59-62, January 1994.
- [Bax94] Walter T. Bax, *A $\Sigma\Delta$ Frequency Discriminator Based Synthesizer*, M. Eng. Dissertation, Carleton University, Ottawa, Canada, December 1994.
- [Bax95] Walt T. Bax et. al., "A $\Sigma\Delta$ Frequency Discriminator Based Synthesizer", *IEEE Proc. ISCAS*, vol. 1, pp. 1-4, Seattle, WA, April 29-May 3, 1995.
- [Bax96] Walt T. Bax et. al., "A Single-Loop Second-Order $\Delta\Sigma$ Frequency Discriminator", *IEEE-CAS Region 8 Proc. Workshop on Analog and Mixed Signal IC Design*, pp. 26-31, Pavia, Italy, September 13-14, 1996.
- [Bax98a] Walt T. Bax and Miles A. Copeland, "A GSM Modulator Using a Delta-Sigma Frequency Discriminator Based Synthesizer" *IEEE Proc. ISCAS*, Monterey, CA, May 31-June 3, 1998.
- [Bax98b] Walt T. Bax et. al., "A GSM Demodulator Based on a Delta-Sigma Frequency Discriminator with Improved Input Sensitivity" *Proc. 24th European Solid-State Circuits Conference*, pp. 264-267, The Hague, Netherlands, September 22-24, 1998.
- [Bear94] R. Douglas Beards and Miles A. Copeland, "An Oversampling Delta-Sigma Frequency Discriminator", *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 41, no. 1, pp. 26-32, January 1994.

-
- [Benn48] W. R. Bennett, "Spectra of Quantized Signals", *Bell Syst. Tech. J.*, vol. 27, pp. 446-472, July 1948.
- [Blan92] Alain Blanchard, *Phase-Locked Loops: Application to Coherent Receiver Design*, Krieger Publishing Company, Malabar, Florida, 1992.
- [Cand74] James C. Candy, "A Use of Limit Cycle Oscillations to Obtain Robust Analog-to-Digital Converters", *IEEE Trans. on Communications*, vol. COM-22, no. 3, pp. 298-305, March 1974.
- [Cand85] James C. Candy, "A Use of Double Integration in Sigma Delta Modulation", *IEEE Trans. on Communications*, vol. COM-33, no. 3, pp. 249-258, March 1985.
- [Couc93] L. W. Couch, *Digital and Analog Communication Systems*, 4th edition, Macmillan, New York, 1993.
- [Craw84] James A. Crawford, "Understanding the Specifics of Sampling in Synthesis: Part I", *Microwaves & RF*, pp. 120-144, August 1984.
- [Craw94] James A. Crawford, *Frequency Synthesizer Design Handbook*, Artech House Inc., 1994.
- [Daup97] P. Leonard Dauphinee, "Balanced 1.5GHz Voltage Controlled Oscillator with Integrated LC Resonator", *Proc. IEEE International Solid-State Circuits Conference*, pp. 390-391, San Francisco, CA, February 6-8, 1997.
- [Demp98] A. G. Dempster and M. D. Macleod, "IIR Digital Filter Design Using Minimum Adder Multiplier Blocks", *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 45, no. 6, pp. 761-763, June 1998.
- [Fenk97] Josef Fenk, "Highly Integrated RF-IC's for GSM and DECT Systems - A Status Review", *IEEE Trans. on Microwave Theory and Techniques*, vol. 45, no. 12, pp. 2531-2539, December 1997.
- [Ferg90] P. F. Ferguson et. al., "One Bit Higher Order Sigma-Delta A/D Converters", *IEEE Proc. ISCAS*, vol. 2, pp. 890-893, New Orleans, LA, May 1-3, 1990.
- [Fili97] Norman M. Filiol et. al., "An Lower ISM Band Frequency Synthesizer and GMSK Data Modulator", *Proc. 23th European Solid-State Circuits Conference*, pp. 60-63, Southampton, U.K., September 16-18, 1997.
- [Gray89] Robert M. Gray, "Spectral Analysis of Quantization Noise in a Single-Loop Sigma-Delta Modulator with dc Input", *IEEE Trans. on Communications*, vol. 37, no. 6, pp. 588-599, June 1989.

-
- [GSM98] European Telecommunications Standard Institute, *Digital cellular telecommunications system (Phase 2+); Radio transmission and reception (GSM 05.05 version 6.2.0 Release 1997)*, TS 100 910, July 1998.
- [Hada91] R. Hadaway et al., "A sub-micron BiCMOS technology for Telecommunications", *Journal of Microelectronics Engineering*, vol 15, pp. 513-516, 1991.
- [HP243] Hewlett Packard, *Fundamentals of the Z-Domain and Mixed Analog/Digital Measurements*, Application Note 243-4.
- [Jens95] J. Jensen et. al., "A 3.2-GHz 2nd Order Delta-Sigma Modulator Implemented in InP HBT technology", *IEEE Journal of Solid-State Circuits*, vol. 30, no. 10, pp. 1119-1127, October 1995.
- [John93] D. A. Johns and D. M. Lewis, "Design and Analysis of Delta-Sigma Based IIR Filters", *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 40, no. 4, pp. 233-240, April 1993.
- [Jou97] S. Jou et. al., "A pipelined Multiplier Accumulator Using a High-Speed, Low-Power Static and Dynamic Full Adder Design", *IEEE Journal of Solid-State Circuits*, vol. 32, no. 1, pp. 114-118, January 1997.
- [Kuo75] B. C. Kuo, *Automatic Control Systems*, 3rd ed., Prentice Hall, Englewood Cliffs, NJ, 1975.
- [Laak98] Timo I. Laakso and V. Valimaki, "Suppression of Transients in Variable Recursive Digital Filters with a Novel and Efficient Cancellation Method", *IEEE Trans. on Signal Processing*, vol. 46, no. 12, pp. 3408-3414, December 1998.
- [Lee87] Wai L. Lee and Charles G. Sodini, "A Topology for Higher Order Interpolative Coders", *IEEE Proc. ISCAS*, vol. 4, pp. 459-462, Philadelphia, PA, May 4-7, 1987.
- [Lu93] F. Lu and H. Samueli, "A 200-MHz CMOS pipelined multiplier-accumulator using a quasi-domino dynamic full-adder cell design", *IEEE Journal of Solid-State Circuits*, vol. 28, no. 2, pp. 123-132, February 1993.
- [Maca91] Raymond C. V. Macario, *Personal & Mobile Radio Systems*, Peter Peregrinus Ltd., London, U.K., 1991.
- [Maki95] Keiko. Makie-Fukuda et. al., "Measurement of Digital Noise in Mixed-Signal Integrated Circuits", *IEEE Journal of Solid-State Circuits*, vol. 30, no. 2, pp. 87-92, February 1995.

-
- [McCl92] M. R. McClure, "Residual Phase Noise of Digital Frequency Dividers", *Microwave Journal*, pp. 124-130, March 1992.
- [Mohi96] Rishi Mohindra, "Isolator for DECT Open loop Modulation", *RF Design Magazine*, pp. 30-42, January 1996.
- [Moto83] Motorola, *MC12040 Phase-Frequency Detector*, data sheet, 1983.
- [Nors97] Steven R. Norsworthy et. al., *Delta-Sigma Data Converters*, IEEE Press, Piscataway, NJ, 1997.
- [Pasu79] S. Pasupathy, "Minimum Shift Keying: A Spectrally Efficient Modulation", *IEEE Communications Magazine*, pp.14-22, July 1979.
- [Perr97] Michael H. Perrott et. al., "A 27-mW CMOS Fractional-N Synthesizer Using Digital Compensation for 2.5-Mb/s GFSK Modulation", *IEEE Journal of Solid-State Circuits*, vol. 32, no. 12, pp. 2048-2060, December 1997.
- [Pete95] Roger L. Peterson, Rodger E. Zeimer and David E. Borth, *Introduction to Spread-Spectrum Communications*, Prentice-Hall Inc., New Jersey, 1995.
- [Rain94] A. J. Rainal, "Eliminating Inductive Noise of External Chip Interconnections", *IEEE Journal of Solid-State Circuits*, vol. 29, no. 2, pp. 126-129, February 1994.
- [Rapp96] Theodore S. Rappaport, *Wireless Communications: Principles & Practice*, Prentice-Hall Inc., New Jersey, 1996.
- [Rile94] Thomas A. D. Riley and Miles A. Copeland, "A Simplified Continuous Phase Modulator Technique", *IEEE Trans. on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 41, no. 5, pp. 321-328, May 1994.
- [Rile98] Thomas A. D. Riley et. al., "A Two-Loop Third-Order Multistage $\Delta\Sigma$ Frequency to Digital Converter", *IEEE Proc. ISCAS*, Monterey, CA, May 31-June 3, 1998.
- [Roul90] David J. Roulston. *Bipolar Semiconductor Devices*, McGraw-Hill, New York, 1990.
- [Stein94] Philip Steiner and Woodward Yang, "Stability Analysis of the Second Order $\Sigma\text{-}\Delta$ Modulator", *IEEE Proc. ISCAS*, vol. 5, pp. 365-368, London, England, May 30-June 2, 1994.
- [Stet95] Trudy Stetzler et. al., "A 2.7-4.5V Single Chip GSM Transceiver RF Integrated Circuit", *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1421-1429, December 1995.

-
- [Su93] David K. Su et. al., "Experimental Results and Modeling Techniques for Substrate Noise in Mixed-Signal Integrated Circuits", *IEEE Journal of Solid-State Circuits*, vol. 28, no. 4, pp. 420-428, April 1993.
- [Wald90] R. H. Walden et. al., "Architectures for Higher-Order Multibit $\Sigma\Delta$ Modulators", *IEEE Proc. ISCAS*, vol. 2, pp. 895-898, New Orleans, LA, May 1-3, 1990.
- [Xion94] F. Xiong, "Modern Techniques in Satellite Communications", *IEEE Communications Magazine*, pp. 84-97, August 1994.
- [Yama97] Taizo Yamawaki et. al., "A 2.7-V GSM RF Transceiver IC", *IEEE Journal of Solid-State Circuits*, vol. 32, no. 12, pp. 2089-2096, December 1997.

Appendix A

GMSK Modulator

Figure A.1: RF and analog section schematic.

Figure A.2: Digital signal processor schematic.

Appendix B

$\Delta\Sigma$ Frequency Discriminator

Figure B.1: BiCMOS $\Delta\Sigma$ FD chip bonding diagram.