Linearisation of RF Power Amplifiers

Nielsen, Per Asbeck; Bruun, Erik; Olsen, Ole; Rebild, Dan

Publication date:
2001

Document Version
Publisher's PDF, also known as Version of record

Citation (APA):
Per Asbeck

Linearisation of RF Power Amplifiers

PhD thesis

Nokia Mobile Phones
and
Department of Information Technology
Technical University of Denmark

March 2001
Linearisation of RF Power Amplifiers

by

Per Asbeck

This thesis is submitted in partial fulfilment of the requirements for the degree: Doctor of Philosophy (Ph.D.)

Ørsted•DTU
Technical University of Denmark

March 2001
Preface

In connection with this Ph.D project I wish to thank several people for their help and kindness.

Professor Erik Bruun (DTU), Ole Olsen (DTU) and Dan Rebild (Nokia Denmark) for supervising this project.
Dan Rebild for starting up this Ph.D. project on behalf of Nokia Denmark A/S and for supporting me in many different ways throughout the project.
Thanks to Ole Olesen for helping me starting and establishing the PhD work and helping me arrange my stay at Stanford University.
Thomas Lee for letting me visits his group at SMIRC (Stanford Microwave Integrated Circuits Laboratory) at Stanford University. I have learned a lot though discussion with the people in the group and through participation in many seminars.
Thanks to Allan Jørgensen (DTU) for CAD support especially with Cadence problems but also in matters of PC, work stations and networks. His help has been indispensable.
Thanks to Nokia Denmark financial support and access to many experienced IC and RF designers.

Note that when I started this project I was affiliated to Institute of Information Technology that later changed it’s name to Ørsted•DTU, Technical University of Denmark.

Thesis written is written in Adobe FrameMaker+SGML 5.5 with 12pt Times New Roman Font. The figures are generated with Visio Technical 5.0+ and Mathcad 2000 Professional. The figures are inserted as objects and windows meta files (WMF) and printed out using Adobe distiller post script driver and Adobe Acrobat 4.0. Many graphs are distinguish though their colour which necessitate a colour printer for the printing.

Per Asbeck
Ørsted•DTU
Technical University of Denmark
Contents

Preface .................................................................................................................................. iii
Contents ................................................................................................................................ v
List of figures ..................................................................................................................... ix
Abstract ................................................................................................................................ xv
Resumé (in danish) ............................................................................................................. xvii
List of Abbreviations ......................................................................................................... xix
List of Notation .................................................................................................................. xxi

Introduction ........................................................................................................................ 1

1 The transmitter .................................................................................................................. 3

1.1 Transmitter architecture ............................................................................................ 3
  1.1.1 Direct conversion ................................................................................................. 3
  1.1.2 Super heterodyne transmitter ............................................................................. 5
  1.1.3 Offset loop transmitter .................................................................................... 6
1.2 Representations of RF signals .................................................................................... 7
  1.2.1 Band limited RF signal ................................................................................... 7
  1.2.2 Complex envelope representation .................................................................... 8
  1.2.3 Filtering of complex envelope signals ............................................................ 9
1.3 Modulation .................................................................................................................. 9
  1.3.1 The modulation process .................................................................................. 11
  1.3.2 IQ modulation ................................................................................................. 11
  1.3.3 QPSK modulation .......................................................................................... 13
  1.3.4 Edge modulation ........................................................................................... 16
  1.3.5 Measuring the signal quality ........................................................................ 18

2 Modelling and Characterising Non-Linearities ............................................................. 21

2.1 Harmonic distortion .................................................................................................. 21
2.2 Intermodulation ......................................................................................................... 23
2.3 Polynomial non-linearities ....................................................................................... 24
  2.3.1 Harmonic distortion ....................................................................................... 25
  2.3.2 Intermodulation ............................................................................................. 26
2.4 Memory-less bandpass non-linearities .................................................................... 28
  2.4.1 Complex envelope bandpass non-linear simulation model ......................... 29
  2.4.2 The limiter non-linearity ............................................................................... 30
  2.4.3 A non-linearity with hysteresis ................................................................. 31

3 The Power Amplifier ...................................................................................................... 33

3.1 Classes of operation ................................................................................................. 34
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.1</td>
<td>Class A</td>
<td>34</td>
</tr>
<tr>
<td>3.1.2</td>
<td>Class B</td>
<td>36</td>
</tr>
<tr>
<td>3.1.3</td>
<td>Classes described by the conduction angle</td>
<td>37</td>
</tr>
<tr>
<td>3.1.4</td>
<td>Class F</td>
<td>42</td>
</tr>
<tr>
<td>3.1.5</td>
<td>Class D</td>
<td>42</td>
</tr>
<tr>
<td>3.1.6</td>
<td>Class E</td>
<td>43</td>
</tr>
<tr>
<td>3.2</td>
<td>Characterisation of the PA</td>
<td>43</td>
</tr>
<tr>
<td>4</td>
<td>Linearisation Techniques</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>The LINC Techniques</td>
<td>46</td>
</tr>
<tr>
<td>4.2</td>
<td>CALLUM</td>
<td>47</td>
</tr>
<tr>
<td>4.3</td>
<td>Cartesian Modulation Feedback</td>
<td>48</td>
</tr>
<tr>
<td>4.4</td>
<td>Polar Modulation Feedback Loop</td>
<td>48</td>
</tr>
<tr>
<td>4.5</td>
<td>Envelope Feedback</td>
<td>50</td>
</tr>
<tr>
<td>4.6</td>
<td>Phase Correcting Feedback</td>
<td>50</td>
</tr>
<tr>
<td>4.7</td>
<td>Feed Forward Linearisation</td>
<td>51</td>
</tr>
<tr>
<td>4.8</td>
<td>Digital Predistortion</td>
<td>51</td>
</tr>
<tr>
<td>4.8.1</td>
<td>Mapping Predistortion</td>
<td>52</td>
</tr>
<tr>
<td>4.8.2</td>
<td>Polar Predistortion</td>
<td>53</td>
</tr>
<tr>
<td>4.8.3</td>
<td>Complex Gain Predistortion</td>
<td>53</td>
</tr>
<tr>
<td>4.8.4</td>
<td>Polynomial predistortion</td>
<td>54</td>
</tr>
<tr>
<td>4.9</td>
<td>Remarks on PA isolation and sensitivity to external disturbance</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>Analysis and design of Linearisation Systems</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>Amplitude linearity of Cartesian Modulation Feedback System</td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td>Polar modulation feedback</td>
<td>60</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Polar linearisation in a digital transmitter</td>
<td>60</td>
</tr>
<tr>
<td>5.2.2.1</td>
<td>Polar loop with an RF-phase shifter</td>
<td>61</td>
</tr>
<tr>
<td>5.2.2.2</td>
<td>Polar loop linearisation using a VCO</td>
<td>62</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Analysis of the Polar loop</td>
<td>63</td>
</tr>
<tr>
<td>5.2.2.1</td>
<td>Loop settlement and linearity compared to Cartesian loop</td>
<td>64</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Dynamics of the polar loop linearisation system</td>
<td>65</td>
</tr>
<tr>
<td>5.2.3.1</td>
<td>Simulating the polar loop</td>
<td>67</td>
</tr>
<tr>
<td>5.2.3.2</td>
<td>Loop gain problem</td>
<td>67</td>
</tr>
<tr>
<td>5.2.3.3</td>
<td>Design guidelines for a envelope feedback system</td>
<td>69</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Gain compensation</td>
<td>69</td>
</tr>
<tr>
<td>5.2.4.1</td>
<td>Simulation of the envelope FB system with gain control</td>
<td>71</td>
</tr>
<tr>
<td>5.2.5</td>
<td>Resume of polar modulation feedback</td>
<td>72</td>
</tr>
<tr>
<td>5.3</td>
<td>Impact of non-linearities on QPSK modulated signal</td>
<td>73</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Phase non-linearity</td>
<td>74</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Amplitude non-linearity</td>
<td>75</td>
</tr>
<tr>
<td>5.3.3</td>
<td>Distortion of a simulated CMOS PA</td>
<td>76</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Summary of impact of non-linearities on QPSK modulated signal</td>
<td>77</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Remarks to the simulations</td>
<td>78</td>
</tr>
<tr>
<td>6</td>
<td>Design of key building blocks</td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction to AM-detectors</td>
<td>81</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Diode AM-detectors</td>
<td>81</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Field Effect Transistor AM-detectors</td>
<td>85</td>
</tr>
<tr>
<td>6.1.3</td>
<td>Linear AM-detectors</td>
<td>85</td>
</tr>
<tr>
<td>6.2</td>
<td>Architectures and design considerations for envelope feed back</td>
<td>86</td>
</tr>
</tbody>
</table>
CONTENTS

6.2.1 The Bandpass filter ................................................................. 87
6.3 Envelope feed back system with non-linear AM-detectors ................. 88
6.4 Experiments with discrete RF components ......................................... 90
   6.4.1 The system ........................................................................ 90
   6.4.2 Characterisation of a commercial PA ........................................ 92
      6.4.2.1 Output versus input power ........................................ 92
      6.4.2.2 Output power controlled by the power control signal ......... 94
      6.4.2.3 Summary of PA measurements ..................................... 95
   6.4.3 Amplitude detection ............................................................ 95
   6.4.4 Circuit boards ...................................................................... 97
   6.4.5 Error amplifier for the feed back loops .................................... 98
   6.4.6 Phase shifter ....................................................................... 100
      6.4.6.1 Measurement of the phase shifter ................................. 101
   6.4.7 Gain Stage ......................................................................... 102
   6.4.8 Closing the power loop ....................................................... 103
   6.4.9 Conclusion on experiments with discrete RF components ........... 104
   6.5 PA design ............................................................................... 105
      6.5.1 The effects of low oxide breakdown voltage ......................... 107
      6.5.2 PA topology ..................................................................... 108
      6.5.3 Designing the output stage .............................................. 109
      6.5.4 Gain stage and interstage matching .................................... 114
      6.5.5 Importance of ground bondwire inductance ......................... 114
      6.5.6 Layout ............................................................................. 115
         6.5.6.1 Inductor design for interstage matching ......................... 116
         6.5.6.2 Capacitor design ....................................................... 118
      6.5.7 Simulations and measurement ............................................ 120
      6.5.8 Conclusion ...................................................................... 121
   6.6 AM-detector design ................................................................... 123
      6.6.1 Passive AM-detector ....................................................... 123
      6.6.2 Design and simulation of the fabricated AM-detector .......... 125
      6.6.3 Measurements ................................................................. 127
   6.7 Active AM-detector ................................................................... 128
      6.7.1 Measurement setup ......................................................... 130
      6.7.2 Comparison between measurement and simulations ............. 134
      6.7.3 Study of the NMOS current mirror ................................... 136
      6.7.4 Performance prediction of AM-detector ......................... 137
   6.8 Linear AM detector .................................................................. 141
   6.9 Resume of chapter 6 ................................................................. 145

Conclusion ......................................................................................... 147

Literature ............................................................................................. 149
Index ..................................................................................................... 153
Appendix A Inphase and quadrature frequency translation
Appendix B Puls shape filter for EDGE
Appendix C Program for on-chip inductor design
Appendix D Publications and patents
List of figures

1.1 Direct up-conversion RF frontend ................................................................. 3
1.2 Double side band (DSB) and single side band (SSB) modulation ............... 5
1.3 Superheterodyne RF transmitter frontend ................................................. 5
1.4 Frequency plan of transmitter with one intermediate frequency .......... 6
1.5 Offset loop transmitter frontend ................................................................. 6
1.6 Generation of the complex envelope signal .............................................. 9
1.7 The digital modulation process and the RF front end ......................... 11
1.8 Single side band modulator (SSB) ............................................................ 12
1.9 IQ-modulation ............................................................................................ 12
1.10 QPSK constellation ................................................................................... 13
1.11 Pulse response of a raised cosine filter ( ) ............................................. 14
1.12 Frequency characteristic of the raised cosine filter .............................. 14
1.13 IQ constellation plot of QPSK modulated random signal
    with raised cosine filter ( ) ................................................................. 15
1.14 Inphase and quadrature signal of QPSK modulated random signal with raised cosine filter ( ) ................................................................. 16
1.15 Edge constellation diagram ...................................................................... 16
1.16 Gauss pulse and linearised Gauss pulse ................................................... 17
1.17 Principle of EDGE modulation definition ............................................. 18
1.18 Measurement of Bit Error Rate (BER) ................................................... 19
1.19 Calculation of the error vector and the error vector magnitude (EVM) ...... 19
1.20 Calculation of adjacent channel power ................................................... 19
2.1 Non-linear system ....................................................................................... 21
2.2 Complex envelope bandpass non-linear simulation model ..................... 29
2.3 Limiter function ( =10, Lin=Lout=1) ....................................................... 30
2.4 A non-linearity with hysteresis ............................................................... 31
3.1 Power amplifier stage ............................................................................... 33
3.2 Power amplifier stage for class A operation ........................................... 34
3.3 The drain current of a transistor in class A operation ............................. 34
3.4 Large signal behaviour of the transistor and the load line ..................... 35
3.5 Class B amplifier ....................................................................................... 36
3.6 Drain current of a transistor in class B operation ..................................... 36
3.7 Drain currents given a class of operation and the conduction ................. 38
3.8 The drain peak current, eq. (3.21), current and DC current, eq. (3.25) normalized by (Vdd=1 and RL=1). All values are for maximum output power ( ) which implies that the AC current is equal to its maximum of one .......................... 40
3.9 Parameters for the drain current, eq. (3.13), the drain peak current, eq. (3.21), and the DC current, eq. (3.25), normalized by (Vdd=1 and RL=1). All values are for maximum
output power ($\alpha=1$) which implies that the AC current is equal to its maximum of one.

3.10 Spectral content of drain current. Left: constant peak drain current. Right: Constant constant value of the fundamental component (constant output power) .................. 41
3.11 Power and efficiency of the output stage as a function of the conduction angle...... 42
3.12 Class F amplifier ................................................................. 43
4.1 Linearisation using the LINC principle .................................................. 46
4.2 CALLUM Linearisation ......................................................................... 47
4.3 Cartesian modulation feedback system .................................................. 48
4.4 Polar modulation feedback ........................................................................ 49
4.5 Polar modulation feedback with a phase shifter ........................................ 49
4.6 Linearisation using the envelope feedback method ..................................... 50
4.7 Phase correcting feedback system ............................................................ 50
4.8 Feed Forward linearisation ...................................................................... 51
4.9 Digital baseband predistortion system ..................................................... 52
4.10 Mapping predistortion (Digital Cartesian Feedback) ................................. 52
4.11 Digital Polar Predistortion ...................................................................... 53
4.12 Digital Complex Gain Predistortion ......................................................... 53
4.13 Isolator avoids disturbance from other transmitters received by the antenna... 54
4.14 A BP-filter in the FB reduces out-of-band disturbances received by the antenna... 54
5.1 RF modulation feedback system ................................................................ 56
5.2 AM/AM and AM/PM characteristic of a PA ............................................. 59
5.3 PA input and output signals inside the feedback system ................................ 59
5.4 Relative error of PA output signal ............................................................. 60
5.5 Traditional polar modulation feedback ..................................................... 61
5.6 Alternative amplitude and phase linearisation ............................................ 61
5.7 Combined frequency synthesis and linearisation ........................................ 62
5.8 Polar linearisation for a digital transmitter ............................................... 63
5.9 Cartesian MFB modified to a envelope feedback system ............................ 64
5.10 Polar loop signal model ........................................................................... 66
5.11 PA output versus the control voltage ....................................................... 68
5.12 ACP versus amplitude delay ................................................................. 69
5.13 Linearisation system DSP controlled gain .............................................. 70
5.14 Circuits for loop gain compensation ....................................................... 70
5.15 Gain compensation function and the resulting PA gain ............................... 71
5.16 RMS amplitude error with and without gain compensation versus maximum amplitude level ................................................................. 71
5.17 Envelope feedback system with gain control .......................................... 72
5.18 Envelope and gain transient simulation .................................................... 72
5.19 Signal flow in receiver and transmitter ..................................................... 73
5.20 Amplitude distortion functions (AM/AM) ............................................... 73
5.21 AM/PM characteristics .......................................................................... 74
5.22 ACP as function of maximum phase error .............................................. 74
5.23 Vector magnitude error as function of maximum phase error ..................... 75
5.24 Third order non-linearity ....................................................................... 75
5.25 ACP versus alpha ................................................................................... 76
5.26 Magnitude vector error versus alpha ....................................................... 76
5.27 ACP, amplitude distortion versus input voltage ........................................ 77
5.28 Relative magnitude error in percent versus input amplitude ....................... 77
5.29 PA input and output .................................................................................. 78
6.37 Gain stage................................................................................................................ 103
6.36 Test bench output from sweeping the control voltage of the phase shifter............. 103
6.35 Test bench output from sweeping the phase between the generators. .................... 102
6.34 Test bench for the phase shifter. ............................................................................. 102
6.33 Photo of the phase shifter........................................................................................ 101
6.32 Aplac simulation of resonator. ................................................................................ 101
6.30 Circuit board for subtraction and amplification in the phase loop.......................... 100
6.29 Circuit board for subtraction and amplification in the power loop........................... 99
6.28 Comparison of the amplitude detectors. ................................................................... 98
6.27 Amplitude detection with a switching mixer. ........................................................... 98
6.26 Amplitude detection with a switching mixer. ........................................................... 97
6.25 Amplitude detector with improved LO drive to the mixer. ...................................... 97
6.24 Amplitude detector characteristic. ............................................................................ 96
6.23 Limiter characteristic. ............................................................................................... 96
6.22 Principle of amplitude detection. .............................................................................. 95
6.21 ACP versus power control voltage............................................................................ 95
6.20 Gain versus control voltage....................................................................................... 94
6.19 Output power versus power control voltage. ............................................................ 94
6.18 ACP measurement..................................................................................................... 93
6.17 Output-power versus input power (WB-CDMA). .................................................... 93
6.16 Generation of the reference signals. ........................................................................... 91
6.15 Linearisation system with commercial available parts. ........................................... 91
6.14 Polar modulation feed back....................................................................................... 90
6.13 Envelope feed back with matched AM-detectors. .................................................... 90
6.12 Linear AM-detector using 2 nonlinear elements and a feed back amplifier............. 85
6.11 Envelope feed back using non-linear AM-detector. ................................................ 89
6.10 Principle of amplitude detection. .............................................................................. 85
6.9 Bandpass filtering using a parallel resonance tank.................................................... 87
6.8 Off-chip and on-chip amplitude detection. ............................................................... 86
6.7 AM detection with biased diode. ............................................................................... 83
6.6 Published FET AM-detectors. (a) and (b) MOSFET MMIC AM-detectors (c) GaAs
AM-detector.......................................................... 84
6.5 Transfer characteristic of diode detector for different loads.................................. 84
6.4 AM-detectors. (a) Positive peak detector, (b) negative peak detector, (c) voltage multiplier, and (d) biased peak detector. ............................................................ 82
6.3 AM detection with biased diode. ............................................................................... 83
6.2 Diode characteristic................................................................................................... 82
6.1 Reference signal and received, aligned signal. .......................................................... 78
5.30 Reference signal and received, aligned signal. .......................................................... 78
5.29 Diode characteristic.................................................................................................. 82
5.28 Amplitude detector characteristic. ........................................................................... 96
5.27 Amplitude detection with improved LO drive to the mixer. ................................... 97
5.26 Linearisation system with commercial available parts. .......................................... 91
5.25 Amplitude detection with high LO signal to the mixer. .......................................... 97
5.24 Comparison of the amplitude detectors. .................................................................. 98
5.23 Limiter characteristic. .............................................................................................. 96
5.22 Principle of amplitude detection. ............................................................................ 95
5.21 Schematic of phase shifter. ..................................................................................... 100
5.20 Aplac simulation of resonator. ................................................................................ 101
5.19 Measurement stage................................................. ............................... 105
5.18 Measurement stage with normalised control voltage.......................... 105
5.17 Measured output power of the gain stage............................................................... 104
5.16 Measurement of the power loop................................................................. 105
5.15 Measurement of output power and ACP sweeping Pin........................................ 93
5.14 Linearisation system with commercial available parts. ........................................ 91
5.13 Generation of the reference signals. ........................................................................ 91
5.12 Polar modulation feed back..................................................................................... 90
5.11 Envelope feed back with matched AM-detectors. ................................................. 90
5.10 Envelope feed back using non-linear AM-detector. ................................................ 89
5.9 Bandpass filtering using a parallel resonance tank................................................... 87
5.8 Off-chip and on-chip amplitude detection. ............................................................... 86
5.7 AM detection with biased diode. ............................................................................... 83
5.6 Published FET AM-detectors. (a) and (b) MOSFET MMIC AM-detectors (c) GaAs
AM-detector.......................................................... 84
5.5 Transfer characteristic of diode detector for different loads.................................. 84
5.4 Amplitude detector characteristic. ........................................................................... 96
5.3 Amplitude detection with improved LO drive to the mixer. ................................... 97
5.2 Principle of amplitude detection. .............................................................................. 85
5.1 Diode characteristic................................................................................................... 82

LIST OF FIGURES

5.30 Reference signal and received, aligned signal. .......................................................... 78
6.1 Diode characteristic................................................................................................... 82
6.2 AM-detectors. (a) Positive peak detector, (b) negative peak detector, (c) voltage multiplier, and (d) biased peak detector. ............................................................ 82
6.3 AM detection with biased diode. ............................................................................... 83
6.4 Transfer characteristic of diode detector for different loads.................................. 84
6.5 Published FET AM-detectors. (a) and (b) MOSFET MMIC AM-detectors (c) GaAs
AM-detector.......................................................... 84
6.6 Principle of amplitude detection. .............................................................................. 85
6.7 Linear AM-detector using 2 nonlinear elements and a feed back amplifier............. 85
6.8 Off-chip and on-chip amplitude detection. ............................................................... 86
6.9 Bandpass filtering using a parallel resonance tank.................................................... 87
6.10 Envelope feed back using non-linear AM-detector. ................................................ 89
6.11 Envelope feed back with matched AM-detectors. .................................................... 90
6.12 Polar modulation feed back..................................................................................... 90
6.13 Generation of the reference signals. ........................................................................ 91
6.14 Linearisation system with commercial available parts. ........................................... 91
6.15 Measurement of output power and ACP sweeping Pin........................................ 93
6.16 Output-power versus input power (WB-CDMA). .................................................... 93
6.17 ACP measurement..................................................................................................... 93
6.18 Output power versus power control voltage. ............................................................ 94
6.19 Gain versus control voltage.................................................................................... 94
6.20 ACP versus power control voltage......................................................................... 95
6.21 Principle of amplitude detection. ............................................................................ 95
6.22 Limiter characteristic. ............................................................................................ 96
6.23 Amplitude detector characteristic. ......................................................................... 96
6.24 Amplitude detector with improved LO drive to the mixer. ................................... 96
6.25 Amplitude detection with high LO signal to the mixer. ........................................ 97
6.26 Amplitude detection with a switching mixer. ........................................................ 97
6.27 Amplitude detection with a switching mixer. ........................................................ 98
6.28 Comparison of the amplitude detectors. .................................................................. 98
6.29 Circuit board for subtraction and amplification in the power loop......................... 99
6.30 Circuit board for subtraction and amplification in the phase loop......................... 100
6.31 Schematic of phase shifter. ..................................................................................... 100
6.32 Aplac simulation of resonator. ................................................................................ 101
6.33 Photo of the phase shifter. ..................................................................................... 101
6.34 Test bench for the phase shifter. ................................................................. 102
6.35 Test bench output from sweeping the phase between the generators.................. 102
6.36 Test bench output from sweeping the control voltage of the phase shifter........ 103
6.37 Gain stage................................................. ............................... 103
6.38 Measurement of gain stage................................................. ............................... 104
6.39 Measured output power of the gain stage............................................................... 104
6.40 Measurement of the power loop................................................................. 105
6.41 PA measurement with power loop................................................................. 105
6.42 PA measurement without power loop............................................................... 106
6.43 Measurements with normalised control voltage................................................ 106
6.44 Power amplifier and load. ..................................................................................... 107
6.45 Two stage power amplifier. ................................................................. 108
6.46 Matching stages (a) L-match (b) T-match (c) Pi-match. ........................................ 109
6.47 Output transistor loaded with a high-Q LC-tank. ................................................ 109
6.48 Simulated normalized output power and PAE versus load resistance

6.49 Class B transistor stage. Drain current and voltage at two drive levels.

Left: Normal linear operation. Right: Drain current bifurcated due to knee voltage.

6.50 Drain currents as function of gate drive level.

6.51 The resonance of the circuit depends on the actual Q of the LC due to the drain-source capacitance (fc=high Q resonance, fc=low Q resonance).

6.52 Matching to the large signal model of the output transistor.

6.53 Resulting transformation resistance as function of the inductance in the 2nd L-match.

6.54 Fabricated PA circuit.

6.55 PA with modelled on-chip and off-chip ground.

6.56 Output power as a function of ground bondwire.

6.57 Layout of the PA.

6.58 MathCad interactive inductor optimization.

6.59 Inductor model with and without shielding to substrate.

6.60 Parallel resonance model of inductor.

6.61 Lateral flux capacitor.

6.62 Measured PA characteristic and PCB photo.

6.63 Photo of entire chip and of the inductor.

6.64 AM-detector. Left: Drain biased. Right: Gate biased.

6.65 NMOS AM-detector. a) drain biased b) gate biased.

6.66 Fabricated amplitude detector.

6.67 PA with on-chip AM-detector.

6.68 Drain biased detector. DC-feed through inductor L1.

6.69 Layout of AM-detector.

6.70 PCB: Left schematic, Right PCB photo.

6.71 Measurement setup for AM-detector.

6.72 Measurement of the AM-detector.

6.73 Amplitude detector with two RF inputs and subtraction circuit.

6.74 Drain current of the detecting transistor as function of RF input amplitude.

6.75 DC currents in AM-detector as function of the RF amplitude.

6.76 Simulation setup of the AM-detector.

6.77 PCB test board with AM-detector.

6.78 Measurement setup.

6.79 Input swept with different bias conditions.

6.80 Signal graphs of the amplitude detector w. and w.o. feed back.

6.81 Measured and compensated transfer functions.

6.82 DC detector characteristic: Aref to Apa.

6.83 Chip photo of die with 2 AM-detectors and a PA.

6.84 Comparison of simulated and measured data.

6.85 Principle schematic used for simulation.

6.86 NMOS current mirror stage w. and w.o. cascodes.

6.87 Simulation with different NMOS current mirrors.

6.88 Simulation with different NMOS current mirrors. Zoom.

6.89 AMAM characteristic of closed loop system based on the characteristic of the AM-detector.

6.90 Error vector magnitude of closed loop system.

6.91 Spectral leakage.

6.92 Principle of linear amplitude detection.
6.93  Schematic of linear amplitude detector .............................................................. 141
6.94  Trans-conductance stages for single ended to differential conversion .......... 142
6.95  The limiter consist of gain stages, limiting amplifiers and stages for single ended to differential conversion ................................................................. 142
6.96  Output voltage of AM-detector and fitted line ................................................ 143
6.97  Chip-photo of the linear AM-detector ............................................................. 143
Abstract

This thesis deals with linearisation techniques of RF power amplifiers (PA), PA design techniques and integration of the necessary building blocks in a CMOS technology. The opening chapters introduces the theory of transmitter architectures, RF-signal representation and the principles of digital modulation. Furthermore different types of power amplifiers, models and measures of non-linearities are presented. A chapter is also devoted to different types of linearisation systems.

The work carried out and described in this thesis can be divided into a more theoretical and system oriented treatment of linearisation systems with focus on polar modulation feedback, and a chip oriented part focusing on integrating of separate building blocks of the system on a chip.

The system oriented part of this thesis deals with analog feedback linearisation systems. The Polar modulation feedback system is compared with the more traditional Cartesian modulation feedback system in terms of loop settlement and dependencies between the feedback signals. A method to calculate the distortion functions of the linearisation system (AM/AM and AM/PM) based on the distortion functions of the power amplifier is presented. Also the polar loop architecture and it’s suitability to modern digital transmitters is discussed. A proposal of an architecture that is suitable for digital transmitters, which means that it has an interface to the digital back-end, defined by low-pass signals in polar form, is presented. Simulation guidelines that utilize properties of the polar loop are presented. Analysis of the envelope feedback loop shows some fundamental limitations of the loop gain and the loop bandwidth due to the varying PA gain. Based on these observation a set of design guidelines for an envelope feedback loop is given. The guidelines consider trade-off between output power and necessary filter bandwidth to guarantee a certain distortion. An analysis of the dynamics of the polar loop shows that the non-linear behaviour of the PA only degrades the stability and the precision of the envelope feedback loop. An extension to the proposed architecture is presented which improves the precision of the envelope feedback system without sacrificing the stability of the loop.

The chip oriented part of this thesis deals with building blocks for polar loop linearisation including the PA itself. A description on experiments on a RF phase shifter and an amplitude detection circuit is given. The purpose is to explore the limitation on commercial available parts and to recognize the challenges in polar loop linearisation. The design of a fabricated CMOS PA is presented. The design is carried out in a standard digital Epi-CMOS which allows integration of circuitry such as the linearisation circuit. The amplifier has the highest output power compared to other published class B power in the same process. The design phase including the on-chip inductor and the lateral flux capacitors is described. The other test chips designed are envelope detectors. Three different detectors were designed. Two non-linear detectors and one linear. The two former AM-detectors have been measured. Based on these measurements the achievable spectral leakage and error vector magnitude were predicted.
Resumé (in danish)

Denne afhandling omhandler linearisering af højfrekvens (HF) effektforstærkere, designteknikker til dimensionering af HF effektforstærker samt integration af enkelte byggeblokke i en CMOS chip teknologi.

De indledende kapitler introducerer senderarkitekturer, signalrepræsentation af modulerede HF-signaler samt modeller og mål for ulinære kredsløb. Derudover præsenteres forskellige arkitekturer til linearisering af HF effektforstærker.

Det udførte arbejde kan deles op i en teoretisk- og systemorienteret del om lineariseringssystemer, og en chip orienteret del, der omhandler integration af vigtige byggeblokke.


Der er blevet designet og fabrikerede tre forskellige amplitude detektorer, hvoraf de to er ulinære og den er lineær. De to første foreligger der målinger på, mens den sidste ikke er blevet leveret af fabrikanten. Baseret på målingerne estimeres støj i nabokanalen (adjacent channel power) samt modulationskvaliteten (error vector magnitude).
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A/D</td>
<td>Analog to digital converter</td>
</tr>
<tr>
<td>ACP</td>
<td>Adjacent channel power</td>
</tr>
<tr>
<td>AM</td>
<td>Amplitude modulation</td>
</tr>
<tr>
<td>BP</td>
<td>Bandpass filter</td>
</tr>
<tr>
<td>CALLUM</td>
<td>Combined analogue locked loop</td>
</tr>
<tr>
<td>CDMA</td>
<td>Code division multiple access</td>
</tr>
<tr>
<td>CDMA2000</td>
<td>Upgrade of cdmaOne (IS95)</td>
</tr>
<tr>
<td>CMOS</td>
<td>Complementary metal-oxide-silicon (field-effect transistor)</td>
</tr>
<tr>
<td>D/A</td>
<td>Digital to analogue converter</td>
</tr>
<tr>
<td>EDGE</td>
<td>Enhanced data rates for GSM evolution</td>
</tr>
<tr>
<td>ETSI</td>
<td>European Telecommunications Standards Institute</td>
</tr>
<tr>
<td>EVM</td>
<td>Error Vector Magnitude</td>
</tr>
<tr>
<td>FB</td>
<td>Feedback</td>
</tr>
<tr>
<td>FDD</td>
<td>Frequency division duplex</td>
</tr>
<tr>
<td>FDMA</td>
<td>Frequency division multiple access</td>
</tr>
<tr>
<td>FET</td>
<td>Field effect transistor</td>
</tr>
<tr>
<td>FM</td>
<td>Frequency modulation</td>
</tr>
<tr>
<td>GMSK</td>
<td>Gaussian filtered minimum shift keying</td>
</tr>
<tr>
<td>GPRS</td>
<td>General packet radio service</td>
</tr>
<tr>
<td>GSM</td>
<td>Global system for mobile communication</td>
</tr>
<tr>
<td>HSCSD</td>
<td>High-speed circuit-switched data</td>
</tr>
<tr>
<td>IQ</td>
<td>Inphase quadrature</td>
</tr>
<tr>
<td>ISI</td>
<td>Inter-symbol interference</td>
</tr>
<tr>
<td>LINC</td>
<td>Linear amplification with non-linear components</td>
</tr>
<tr>
<td>LO</td>
<td>Local oscillator</td>
</tr>
<tr>
<td>LP</td>
<td>Lowpass filter</td>
</tr>
<tr>
<td>MOSFET</td>
<td>Metal-oxide-silicon field-effect transistor</td>
</tr>
<tr>
<td>MSK</td>
<td>Minimum shift keying</td>
</tr>
<tr>
<td>PA</td>
<td>Power amplifier</td>
</tr>
<tr>
<td>PD</td>
<td>Phase detector</td>
</tr>
<tr>
<td>PDC</td>
<td>Personal digital cellular</td>
</tr>
<tr>
<td>PSK</td>
<td>Phase-shift keying</td>
</tr>
<tr>
<td>QAM</td>
<td>Quadrature amplitude modulation</td>
</tr>
<tr>
<td>QPSK</td>
<td>Quadrature phase shift keying</td>
</tr>
<tr>
<td>RC</td>
<td>Raised cosine</td>
</tr>
<tr>
<td>RCC</td>
<td>Root raised cosine</td>
</tr>
<tr>
<td>RF</td>
<td>Radio frequencies</td>
</tr>
<tr>
<td>RMS</td>
<td>Root mean square</td>
</tr>
<tr>
<td>TDD</td>
<td>Time division multiple access</td>
</tr>
<tr>
<td>TDMA</td>
<td>Time-division multiple-access</td>
</tr>
<tr>
<td>UMTS</td>
<td>Universal mobile telecom system</td>
</tr>
<tr>
<td>UWC-136</td>
<td>Universal wireless communication</td>
</tr>
<tr>
<td>VCO</td>
<td>Voltage controlled oscillator</td>
</tr>
<tr>
<td>VGA</td>
<td>Variable gain amplifier</td>
</tr>
<tr>
<td>WB-CDMA</td>
<td>Wide-band code division multiple access</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS
List of Notation

\( k \)  
Boltzmann constant or filter gain for a specific frequency

\( R_{50} \)  
50Ω resistance

\( \alpha \)  
amplitude drive level, gain constant, attenuator constant  
or bandwidth parameter for the cosine filter

\( A \)  
amplitude function

\( A(t) \)  
amplitude of the carrier signal

\( a_1, a_2, \ldots, a_i \)  
coefficients in a Fourier expansion

\( A_{1dB} \)  
1-dB compression point

\( AM/AM_{sys}(A) \)  
system amplitude characteristic

\( AM/PM_{sys}(A) \)  
system phase characteristic

\( A_{PA} \)  
amplitude of PA output

\( \arg(z) \)  
argument of complex number

\( \text{asin}(x) \)  
inverse function of sine

\( \beta \)  
gain of amplifier

\( B \)  
gain of amplitude detector

\( b_1, b_2, \ldots, b_i \)  
coefficients in a Fourier expansion

\( Bessel \)  
Bessel function

\( BW \)  
bandwidth of the modulated signal

\( C_{ds} \)  
drain-source capacitance

\( C_{\text{flux}} \)  
capacitance

\( C_p \)  
parameter for an equivalent circuit of an inductor

\( C_{\text{plate}} \)  
capacitance

\( \delta \)  
skin depth

\( d_{avg} \)  
average diameter

\( d_{in} \)  
inner diameter

\( d_{out} \)  
outer diameter

\( \varepsilon \)  
dielectric constant

\( e_1 \)  
error signal

\( f(A) \)  
amplitude distortion function (AM/AM)

\( f_0 \)  
reference frequency

\( f_0, f_{\omega_1}, f_{\omega_2} \)  
resonance frequencies

\( f_c \)  
carrier frequency

\( f_{CL} \)  
closed loop pole

\( f_{CL,\text{min}} \)  
minimum closed loop pole

\( f_{\text{max}} \)  
maximum frequency of amplitude signal

\( f_{\text{mod}} \)  
bandwidth of modulated signal

\( f_{SR} \)  
self resonance frequency

\( G \)  
small signal gain

\( g(A) \)  
phase distortion function (AM/PM)

\( G_0 \)  
DC gain of filter, \( G_F(0) \)

\( g_2 \)  
gain constant
<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_c[t,A(t)]$</td>
<td>loop gain control signal</td>
</tr>
<tr>
<td>$G_F(s)$</td>
<td>filter in amplitude loop</td>
</tr>
<tr>
<td>$G_{lin}$</td>
<td>small signal gain</td>
</tr>
<tr>
<td>$G_{PA}$</td>
<td>PA gain</td>
</tr>
<tr>
<td>$\eta$</td>
<td>power efficiency</td>
</tr>
<tr>
<td>$HD_i$</td>
<td>$i$-th harmonic distortion point</td>
</tr>
<tr>
<td>$\eta_{drain}$</td>
<td>drain power efficiency</td>
</tr>
<tr>
<td>$Hilb[f(t)]$</td>
<td>Hilbert transform of $f(t)$</td>
</tr>
<tr>
<td>$H_{RC}(f)$</td>
<td>raised cosine filter</td>
</tr>
<tr>
<td>$H_{rec}(f)$</td>
<td>receive filter</td>
</tr>
<tr>
<td>$H_{RRC}(f)$</td>
<td>root raised cosine filter</td>
</tr>
<tr>
<td>$I(t)$</td>
<td>inphase signal</td>
</tr>
<tr>
<td>$I_d$</td>
<td>drain current</td>
</tr>
<tr>
<td>$I_{diode}$</td>
<td>diode current</td>
</tr>
<tr>
<td>$I_{dQ}$</td>
<td>offset quiescent current</td>
</tr>
<tr>
<td>$Im[z]$</td>
<td>imaginary part of $z$</td>
</tr>
<tr>
<td>$IM_i$</td>
<td>$i$-th intermodulation components</td>
</tr>
<tr>
<td>$I_{out}$</td>
<td>output Current</td>
</tr>
<tr>
<td>$I_p$</td>
<td>amplitude of drain current</td>
</tr>
<tr>
<td>$IP_{3i}$</td>
<td>3rd order intermodulation intercept point</td>
</tr>
<tr>
<td>$IP_{th}$</td>
<td>harmonic distortion intercept point</td>
</tr>
<tr>
<td>$I_Q$</td>
<td>PA quiescent current</td>
</tr>
<tr>
<td>$I_s$</td>
<td>saturation current</td>
</tr>
<tr>
<td>$\phi(t)$</td>
<td>phase of the carrier signal</td>
</tr>
<tr>
<td>$len$</td>
<td>length of inductor</td>
</tr>
<tr>
<td>$L_{in}$</td>
<td>input parameter of a limiter</td>
</tr>
<tr>
<td>$L_{out}$</td>
<td>output parameter of a limiter</td>
</tr>
<tr>
<td>$L_p$</td>
<td>parameter for an equivalent circuit of an inductor</td>
</tr>
<tr>
<td>$L_{swall}$</td>
<td>length of side wall</td>
</tr>
<tr>
<td>$m$</td>
<td>modulation index</td>
</tr>
<tr>
<td>$m$</td>
<td>transformation factor</td>
</tr>
<tr>
<td>$n$</td>
<td>emission coefficient</td>
</tr>
<tr>
<td>$N$</td>
<td>number of turns in an inductor</td>
</tr>
<tr>
<td>$p$</td>
<td>sign of error signal</td>
</tr>
<tr>
<td>$P_{ac}$</td>
<td>AC power delivered to the load</td>
</tr>
<tr>
<td>$P_{DC}$</td>
<td>DC power consumption</td>
</tr>
<tr>
<td>$P_{in}$</td>
<td>input power</td>
</tr>
<tr>
<td>$P_{inc}$</td>
<td>input power to diode detector</td>
</tr>
<tr>
<td>$P_{out}$</td>
<td>output power</td>
</tr>
<tr>
<td>$P_{tap}$</td>
<td>power consumed by the amplitude detector</td>
</tr>
<tr>
<td>$q$</td>
<td>elementary charge</td>
</tr>
<tr>
<td>$\Theta$</td>
<td>phase</td>
</tr>
<tr>
<td>$\Theta(t)$</td>
<td>phase of the carrier signal</td>
</tr>
<tr>
<td>$Q(t)$</td>
<td>quadrature signal</td>
</tr>
<tr>
<td>$Q1,Q2$</td>
<td>quality factor</td>
</tr>
<tr>
<td>$\Theta_c$</td>
<td>conduction angle</td>
</tr>
<tr>
<td>Notation</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
</tr>
<tr>
<td>$Q_L$</td>
<td>inductor quality value</td>
</tr>
<tr>
<td>$Re[z]$</td>
<td>real part of $z$</td>
</tr>
<tr>
<td>$R_L$</td>
<td>PA load resistance</td>
</tr>
<tr>
<td>$R_p$</td>
<td>parameter for an equivalent circuit of an inductor</td>
</tr>
<tr>
<td>$R_s$</td>
<td>series resistance of inductor</td>
</tr>
<tr>
<td>$\hat{s}(t)$</td>
<td>complex envelope of carrier signal</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>resistivity</td>
</tr>
<tr>
<td>$s(t)$</td>
<td>carrier signal</td>
</tr>
<tr>
<td>$S$</td>
<td>RF input signal to a linearisation system</td>
</tr>
<tr>
<td>$s_1, s_2, s_3, \ldots, s_k$</td>
<td>complex symbols used for modulation</td>
</tr>
<tr>
<td>$s_2$</td>
<td>downconverted feedback signal</td>
</tr>
<tr>
<td>$S_{in}$</td>
<td>RF input signal</td>
</tr>
<tr>
<td>$S_{out}$</td>
<td>RF output signal</td>
</tr>
<tr>
<td>$sp$</td>
<td>spacing of metals in a inductor</td>
</tr>
<tr>
<td>$S_{PA}$</td>
<td>RF PA output signal</td>
</tr>
<tr>
<td>$S_{ref}$</td>
<td>RF reference signal</td>
</tr>
<tr>
<td>$\tau$</td>
<td>group delay or time constant</td>
</tr>
<tr>
<td>$T_0$</td>
<td>reference temperature</td>
</tr>
<tr>
<td>$THD$</td>
<td>total harmonic distortion</td>
</tr>
<tr>
<td>$\tau_{max}$</td>
<td>maximum group delay</td>
</tr>
<tr>
<td>$t_{metal}$</td>
<td>metal height</td>
</tr>
<tr>
<td>$T_s$</td>
<td>symbol period</td>
</tr>
<tr>
<td>$u(t)$</td>
<td>feedback amplitude</td>
</tr>
<tr>
<td>$V_{bias}$</td>
<td>bias voltage</td>
</tr>
<tr>
<td>$V_{dd}$</td>
<td>supply voltage</td>
</tr>
<tr>
<td>$V_{diode}$</td>
<td>diode voltage</td>
</tr>
<tr>
<td>$V_{max}$</td>
<td>maximum drain voltage</td>
</tr>
<tr>
<td>$V_{out}$</td>
<td>output voltage</td>
</tr>
<tr>
<td>$V_{PA}$</td>
<td>PA amplitude</td>
</tr>
<tr>
<td>$V_{PC}(t)$</td>
<td>power control signal</td>
</tr>
<tr>
<td>$V_{ref}$</td>
<td>reference amplitude</td>
</tr>
<tr>
<td>$V_{sat}$</td>
<td>saturation voltage</td>
</tr>
<tr>
<td>$W$</td>
<td>width of inductor</td>
</tr>
<tr>
<td>$\omega_0$</td>
<td>Reference frequency for complex envelope</td>
</tr>
<tr>
<td>$\omega_1, \omega_2$</td>
<td>frequencies of carriers</td>
</tr>
<tr>
<td>$x(t), y(t)$</td>
<td>input and output of a non-linear system</td>
</tr>
<tr>
<td>$X_{Cds}$</td>
<td>reactance of capacitor</td>
</tr>
<tr>
<td>$Z_0$</td>
<td>the set of all integers</td>
</tr>
<tr>
<td>$Z_t$</td>
<td>tank impedance</td>
</tr>
</tbody>
</table>
Introduction

This thesis deals with linearisation techniques of RF power amplifiers (PA), PA design techniques and integration of the necessary building blocks in a CMOS technology. Linearisation enables the utilization of non-linear and power efficient power amplifiers for amplitude modulated signal. A linearisation system in a cheap CMOS technology is a cost and power efficient alternative to linear power amplifiers, or power amplifiers using power back-off. Below the motivation is explained in more details. Thereafter a survey of this thesis is elaborated.

Background and motivation

Most wireless communication today is based on digital coded signals. Even voice and motion picture are converted to digital information before transmission. There are several reasons for this. Digital information can be compressed (redundancy removed). It can be protected against artifacts in the transmission-channel (error detection and correction) and trade-offs between quality and bandwidth (voice and picture) can easily be carried out compared to analog systems.

The availability of communication channels (frequency bands and cables) are limited and costly which motivates bandwidth efficient modulation of signals. This is accomplished using amplitude and phase modulated carrier signals which are difficult to amplify in a power efficient way before they are sent into the channel (antenna or cable). The reason being that power amplifiers are most power efficient near the maximum output level and that the efficiency drops fast for decreasing signal levels. Furthermore they are non-linear which result in a signal quality degradation and disturbance of neighbouring channels.

One solution to avoid signal degradation and disturbances is to build a power amplifier capable of handling large signals power-efficiently, and then use it only to output signal-levels at a fraction of it ability (power back-off).

Another solution is to employ linearisation to compensate for the non-linearity of the power amplifier. The linearisation system either modifies the signal in front of the PA or it adds a signal to the signal coming out of the PA. Either way, the goal is to make it look like the signal never experienced a non-linearity. In this way it is possible to use the PA close to its highest output level which leads to a more power efficient amplifier. The power consumed by the circuitry used for the linearisation must be taken into consideration when the overall power efficiency of the system is determined.

Linearisation systems are designed to interface with the input and the output signal of the power amplifiers. They are often build next to the power amplifier in the portable device where they occupy a significantly amount of space. Furthermore they are costly due to the number of parts used in the system. The perspective, and one of the goals of integrating the linearisation system, or parts of it, on a chip is to reduce cost and occupied space. Depending on the practical requirements, such as radio specifications and costs, it would be beneficial to integrate the PA with the linearisation system. In this way new portable devices would look like the old ones but provide transmission of bandwidth efficient modulated signals in a power efficient way. (continues)
INTRODUCTION

Thesis contents description
This thesis contains six chapters. The last two major chapters describes the work carried out throughout this Ph.D. project. The first four chapters cover the theory necessary to understand systems in terms of architecture, signal quality and power efficiency.

Chapter 1 gives an introduction to transmitter architecture including direct conversion and architectures with one intermediate frequency. An efficient way of representing modulated RF signals called complex envelope is described which is useful for circuit simulators and hand calculations. The last part of the chapter describes the principle of digital modulation and gives two examples (QPSK and EDGE).

Chapter 2 describes the models and measures of non-linearities.

Chapter 3 introduces the concept of power amplifiers and describes the relation between conduction angle, output power and power efficiency.

Chapter 4 is an introduction to linearisation systems. Several types of digital and analog schemes with one or two power amplifiers are described.

Chapter 5 deals with analog feedback linearisation systems, in particular polar envelope modulation feedback. This system is compared with the more traditional Cartesian modulation feedback system in terms of loop settlement and dependencies between the feedback signals. A method to calculate the distortion functions of the linearisation system (AM/AMsys and AM/PMsys) based on the distortion functions of the power amplifier (AM/AMPA and AM/PMPA) is presented. Also the polar loop architecture and its suitability to modern digital transmitters is discussed. A proposal to an architecture that is suitable for digital transmitters which means that it has an interface to the digital back-end defined by low-pass signals in polar form is presented. Parts of the hardware can be eliminated from the traditional polar modulation feedback loop. It is found that the nature of the polar loop offers some possibilities for reducing the simulation complexity. This is utilized by a set of simulation guidelines presented in the chapter. Analysis of envelope feedback loop of a first order system showed some basic limitations of loop gain and loop bandwidth due to the varying PA gain. Based on these observation a set of design guidelines for an envelope feedback loop is given. The guidelines consider trade-off between output power and necessary filter bandwidth to guarantee a certain maximum distortion. An analysis of the dynamics of the polar loop also shows that the non-linear behaviour of the PA degrades the stability and the precision of the envelope feedback loop. An extension to the proposed architecture is presented which improves the precision of the envelope feedback loop without sacrificing the stability.

Chapter 6 deals with building blocks for polar loop linearisation including the PA itself. The chapter begins with a description on experiments with a phase shifter and an amplitude detection circuit. The purpose is to explore the limitation on commercial available parts and to recognize the challenges in polar loop linearisation. Then the design of a fabricated CMOS PA is presented. The design is carried out in a standard digital Epi-CMOS since this is a cost-efficient process that allows integration of other circuitry such as the linearisation circuit. The amplifier has the highest output power compared to other published class B power in the same process. The design phase including the on-chip inductor and the lateral flux capacitors is described. The other test chips designed are envelope detectors. Since there have been few publications on envelope detection, especially in CMOS, three different detectors were designed during this project. Two non-linear detectors and one linear. The latter was unfortunately not available at the time of the writing due to a delay at the foundry therefore measurements are not presented. The other two detectors have been measured. Based on these measurements of linearity and sensitivity the achievable spectral leakage for a EDGE signal is predicted.
Chapter 1

The transmitter

1.1 Transmitter architecture

In this chapter we will look at some traditional transmitter architectures. The receiver part which is present in all wireless devices will not be treated in this work because the focus will be on power amplifiers and linearisation which are parts of the transmitter. The purpose of modern transmitters is to convert the digital information to a signal suitable for transmission over the air. In order to choose the right architecture several aspects have to be taken into consideration. Some of them are the noise properties, linearity and power consumption of the building blocks. Also what kind of environment the transmitter is going to operate in is important. Specification such as allowable emission of power outside the transmission frequency band, dynamic range and power control will have a big impact and the choice of architecture.

1.1.1 Direct conversion

A block diagram of a transmitter is shown in Figure 1.1. This type of transmitter is called a direct conversion transmitter because the modulated signal is translated directly up to the carrier frequency. The baseband processor represents the application and the digital modulator. The application could for instance be parts of a mobile phone controlling the microphone, loudspeaker, the keyboard and the display. The data that the application generates, are modulated in the digital modulator to bring it on a form suitable for transmission over the air. Almost any modulated

![Figure 1.1: Direct up-conversion RF frontend.](image-url)
RF signal can be described by two functions, an inphase and a quadrature signal [1]. These signal are called \( I(t) \) and \( Q(t) \) in Figure 1.1.

The RF frontend consist of an local oscillator (LO), a -90 degree phase shifter, two mixers, a combiner (adder or subtractor) and a power amplifier (PA). The LO is normally a frequency synthesizer with a programmable output frequency corresponding to the different RF channels used in the radio system. The D/A converters converts the digital signal to a time continuous, analog signal. These signals are multiplied by two RF carrier signals 90 degrees out of phase. The carriers are in quadrature. The multiplication is carried out by the mixer. The multiplication corresponds to a frequency shift of the signals \( I(t) \) and \( Q(t) \). The output of the mixers are either subtracted or added to obtain the RF signal to be transmitted. The RF signal is amplified by the power amplifier before it is fed to the antenna. The resulting output signal of the transmitter is a amplified version of the signal:

\[
s(t) = I(t) \cdot \cos(\omega_c t) - Q(t) \cdot \sin(\omega_c t)
\]  

(1.1)

The \( I(t) \) and \( Q(t) \) signals are directly related to an amplitude and phase modulated carrier signal through the following relation.

\[
A(t) \cos(\omega_c t + \phi(t)) = I(t) \cdot \cos(\omega_c t) - Q(t) \cdot \sin(\omega_c t).
\]

(1.2)

where

\[
I(t) = A(t) \cos \phi(t)
\]

\[
Q(t) = A(t) \sin \phi(t)
\]

The information, \( I(t) \) and \( Q(t) \), contained in the RF signal can be extracted at the receiver by multiplying one of the carrier signals (\( \cos(\omega_c t) \) or \( \sin(\omega_c t) \)) to the received signal. If we multiply the transmitted signal, \( s(t) \), by the carrier signal \( \cos(\omega_c t) \) we obtain a low frequency component corresponding to the inphase signal.

\[
s(t) \cos(\omega t) = \frac{1}{2} [I(t) + I(t) \cos(2\omega t) - Q(t) \sin(2\omega t)]
\]

(1.3)

\[
LowPass[s(t) \cdot \cos(\omega t)] = \frac{1}{2} I(t)
\]

(1.4)

Similarly we can obtain the quadrature component by multiplying the sent signal, \( s(t) \), by the quadrature component of the local oscillator.

\[
LowPass[s(t) \cdot \sin(\omega t)] = -\frac{1}{2} Q(t)
\]

(1.5)

The recovered inphase and quadrature signal can be fed to a digital demodulator and the data sent by the application can be derived.

One obvious reason for combining the two modulated information signals is that they occupy the same frequency spectrum and still it is possible to separate the information again. So in this way spectral efficiency is achieved. Spectral efficiency has been an important topics since the early radio days where single side band (SSB) modulation [2] was important due to the fact that half the power could be saved (see spectrum in Figure 1.2). The present architecture in Figure 1.1 is also able to accommodate SSB provided that the modulator delivers the Hilbert transformed of the other signal. E.g.

\[
I(t) = Hilbert[Q(t)]
\]

(1.6)
where the Hilbert transform [3] is defined as follows:

$$\text{Hilbert}[x(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t - \tau} d\tau$$  \hspace{1cm} (1.7)

The transform can be seen as a 90 degree phase shifter for all frequencies.

**Figure 1.2:** Double side band (DSB) and single side band (SSB) modulation.

Single side band modulated signals only occupies half the bandwidth compared to the case where \(I(t)\) and \(Q(t)\) are unrelated, but notice that only half the information is propagated \(I(t)\). So IQ-modulation with independent inphase and quadrature signal is as spectral efficient (information per bandwidth) as SSB. SSB probably became uninteresting as speech was coded digitally. Then it is always possible to split the incoming signal into two digital signals just by picking every second sample alternately. Thereby the Hilbert transform which requires extra processing power is avoided.

### 1.1.2 Super heterodyne transmitter

The superheterodyne transmitter in Figure 1.3 is characterized by having one intermediate frequency before translating up to the carrier frequency. The advantage is that some of the unwanted frequency components originating from non-linearities in e.g. the mixers can be removed by filters. Due to the application of an intermediate frequency two local oscillators and two mixer stages are need. The first mixer stage is an inphase and quadrature mixer (IQ-mixer) similar to the one used in the direct conversion transmitter in the previous section (Section 1.1.1).

It is seen that the IQ-mixer stage in Figure 1.3 can be perceived as one stage if a complex number formulation is used. The signal \(s_f\) at the intermediate frequency can be described as follows.

$$s_1(t) = \text{Re}[(I(t) + j \cdot Q(t)) \cdot e^{j \cdot 2\pi \cdot f_1}]$$

$$= I(t) \cdot \cos (2\pi \cdot f_1) - Q(t) \cdot \sin (2\pi \cdot f_1)$$  \hspace{1cm} (1.8)
The inphase signal, $I(t)$, and quadrature signal, $Q(t)$, are perceived as one complex signal being multiplied by a complex exponential function. Only one multiplication is carried out which can be seen as one mixer stage.

The intermediate frequency is determined by the first LO frequency, $f_1$. Two frequencies are generated by the second mixer stage $f_1 + f_2$ and $f_1 - f_2$. The bandpass filter after the second mixer (BP1) selects one of the frequencies and reject the other one. The unwanted frequency is called the image frequency. Either of the two possible frequencies must be chosen for the pass-band of the bandpass filter BP1.

$$f_c = f_{\text{pass}} = f_1 + f_2$$  \hspace{1cm} (1.9)

$$f_c = f_{\text{pass}} = f_1 - f_2$$  \hspace{1cm} (1.10)

The first bandpass filter, BP2, is only necessary if some of the blocks preceding the filter are non-linear or noisy. In this case harmonics of the LO or noise outside the modulation band can be eliminated. This is a way to full fill requirements of a radio system using poor (non-linear or noisy) building blocks. Of course only out of band specifications are improved. It is not possible to improve inband distortion or noise, but fortunately it turns out that out of band requirements are the toughest (see Chapter 5.3).

### 1.1.3 Offset loop transmitter

The offset loop transmitter (Figure 1.5) is also using an intermediate frequency. Opposed to a super heterodyne transmitter the second LO signal is used inside a phase locked loop (PLL) to translate the carrier down to the intermediate frequency. The transmitter consist of a direct conversion transmitter as the first stage. The second stage, the PLL, consist of a phase detector
(PD), a low pass filter (LP), a voltage controlled oscillator (VCO), a down-conversion mixer and a bandpass filter (BP). The bandpass filter is used to reject the high frequency image. The carrier is generated by a voltage controlled oscillator (VCO) inside the PLL. As the carrier signal is generated from a VCO it is obvious that it contains no envelope variations. Therefore this architecture only accommodate constant envelope modulation.

The advantage of this architecture is that the bandwidth of the PLL can be set so low as the bandwidth of the modulated signal\(^1\). The modulated signal will propagate though the PLL but noise outside the modulation bandwidth originates only from the VCO and the lowpass filter. Any other out of modulation-band noise from the blocks in the PLL are removed by the lowpass filter. The channel selection is normally made by changing the frequency of the second LO, \(f_2\), but in principle the first LO can be used too. The carrier frequency is equal to the VCO frequency which is the sum of the two oscillator frequencies.

\[
f_c = f_{VCO} = f_1 + f_2
\]  

(1.11)

**1.2 Representations of RF signals**

Analysing and developing RF architecture systems requires a compact and easy way to mathematically represent signals. Below different RF signal representations and important properties are described. The representations can be used to simulate non-linear, multiple IF systems very efficiently.

**1.2.1 Band limited RF signal**

Carrier signals used in communications systems are typical bandpass signals which means that the signal bandwidth is much smaller than the carrier frequency. This also implies that the phase and the amplitude of the carrier signal changes slowly compared to the period of the carrier frequency. It is therefore more efficient to describe these functions instead of the carrier signal itself. Two representations are possible depending if polar or cartesian coordinates are chosen. A band limited phase and amplitude modulated RF signal can be represented in two of following ways:

1) By a amplitude and phase function

\[
s(t) = A(t) \cos(\omega_c t + \varphi(t))
\]  

(1.12)

where

- \(A(t)\) is the amplitude modulation function
- \(\varphi(t)\) is the phase modulation function
- \(\omega_c\) is the carrier frequency \(\omega_c = 2\pi \cdot f_c\)

2) Or by a inphase and quadrature component

\[
s(t) = I(t) \cos(\omega t) - Q(t) \sin(\omega t)
\]  

(1.13)

where

- \(I(t)\) is the inphase component
- \(Q(t)\) is the quadrature component

---

1. in the following called the modulation bandwidth
It is noticed that \( s(t) \) in equation (1.13) can be perceived as the output of an IQ-up-conversion mixer with \( I(t) \) and \( Q(t) \) as input signals. If \( s(t) \) is band limited with the bandwidth, \( BW \), much smaller than the carrier frequency, \( f_c \), then \( A(t), \phi(t), I(t) \) and \( Q(t) \) are slowly varying functions relative to the period of the carrier frequency, \( 1/f_c \). The relation between \( A(t), \phi(t), I(t) \) and \( Q(t) \) can be derived through complex numbers representation.

\[
\begin{align*}
    s(t) &= Re[(I(t) + j \cdot Q(t)) \cdot e^{j\omega_0 t}] \\
    &= Re[A(t)e^{j\phi(t)} \cdot e^{j\omega_0 t}]
\end{align*}
\]

which implies the following relation

\[
A(t)e^{j\phi(t)} = I(t) + j \cdot Q(t)
\] (1.15)

Separating the real and the imaginary part one achieve formulas for \( I(t) \) and \( Q(t) \).

\[
\begin{align*}
    Re[A(t)e^{j\phi(t)}] &= A(t)\cos(\phi(t)) = I(t) \\
    Im[A(t)e^{j\phi(t)}] &= A(t)\sin(\phi(t)) = Q(t)
\end{align*}
\]

(1.16) 
(1.17)

Dividing the above equations an expression for the phase \( \phi(t) \) is achieved.

\[
\tan(\phi(t)) = \frac{Q(t)}{I(t)}
\] (1.18)

Squaring and adding the equation and applying the square root the magnitude of \( A(t) \) is achieved

\[
|A(t)| = \sqrt{[I(t)]^2 + [Q(t)]^2}
\] (1.19)

### 1.2.2 Complex envelope representation

The complex envelope representation is a low frequency representation convenient for RF simulation and hand calculations purposes. It can be described by the low pass functions in Section 1.2.1. Formally the complex envelope, \( \tilde{s}(t) \), of an RF signal \( s(t) \) is defined in [33] as the pre-envelope (analytical signal) multiplied by a complex modulation function \( e^{-j\omega_0 t} \)

\[
\tilde{s}(t) = (s(t) + j \cdot Hilb(s(t))) \cdot e^{-j\omega_0 t}
\] (1.20)

where

- \( s(t) \) is real valued RF-signal to be used to calculate the complex envelope
- \( Hilb(*) \) is the Hilbert transform defined in equation (1.7)
- \( \omega_0 \) is a reference frequency

A necessary condition for transformation to hold is that the bandwidth is less than the carrier frequency.

\[
BW < \frac{\omega_0}{2\pi}
\] (1.21)

Adding the Hilbert transform to the original signal in equation (1.20) set the negatives spectral components of the signal to zero. Afterwards the signal spectrum is single sided translated towards DC with the amount of a reference frequency \( \omega_0 \). The necessary steps to calculate the complex envelope are illustrated in Figure 1.6.

A complete representation of a RF signal consist of besides of the complex envelope signal of the reference frequency, \( \omega_0 \), used in the complex modulation function in equation (1.20). In the
typical case the reference frequency is chosen to equal to the carrier frequency, $\omega_0 = \omega_c$. In this case the complex envelope is as follows.

$$\tilde{s}(t) = x_i(t) + j \cdot x_Q(t) = A(t) \cdot e^{j\Psi(t)} \quad (1.22)$$

The original signal can always be obtained by multiplying the complex envelope with the complex exponential function and taking the real part as shown in the equation below.

$$s(t) = Re[\tilde{s}(t)e^{j\omega_0 t}] = I(t) \cdot \cos(\omega t) - Q(t) \cdot \sin(\omega t) \quad (1.23)$$

### 1.2.3 Filtering of complex envelope signals

Filtering of RF signal can be carried out using complex envelop signals in the following way. Assume that we have a bandpass filter with the impulse response $h(t)$ and an input signal $s(t)$ then the complex envelope of the output, $y(t)$, is calculated as the convolution of complex envelope of the impulse response and the complex envelope of the inputs signal divided by two [33].

$$\tilde{y}(t) = \frac{1}{2} (\tilde{h} \cdot \tilde{s}) = \frac{1}{2} \int_{-\infty}^{\infty} \tilde{h}(t - \tau) \tilde{s}(\tau) d\tau \quad (1.24)$$

where $\tilde{h}(t), \tilde{s}(t), \tilde{y}(t)$ are the complex envelope of the signals $h(t), s(t), y(t)$ respectively. Notice that in [33] $h(t)$ is defined as $(1/2)(h(t) + j \cdot Hilb(h(t))) \cdot e^{-j\omega_0 t}$ to avoid the factor $1/2$ in the convolution term.

The advantage of complex envelope representation is seen here because the convolution is carried out on the slow varying complex envelope signal instead of the carrier signal. The same is the case for the bandpass impulse response. If it has a narrow bandwidth compared with the centre frequency of the BP-filter it will be a slow varying function too.

### 1.3 Modulation

In the following some modulation techniques for the most common digital wireless phone systems will be presented. In a phone system many aspects has to be taken into consideration. The most important are spectral efficiency, power consumption and transceiver complexity. Spectral efficiency is as measure for the information packed into a certain bandwidth. It tells how many bits per second are transmitted using a certain bandwidth [bits/s/Hz]. This is important for a phone operator because it limits the maximum number of simultaneous phone call given a certain allocated bandwidth.
The power consumption is also influenced by the type of modulation. Typically the type of filtering (pulse shaping) and the number of symbols in the modulation have a big impact the consumed power. If a filter has an overshoot in the step response it normally leads to higher power consumption. A high filter bandwidth will reduce the spectral efficiency. More symbols in a modulation yields a high spectral efficiency but increases the sensitive to noise. So a higher transmit power has to be used to maintain the signal quality.

Transceiver complexity is important due to cost. The more complicated the modulation gets in terms of filtering and symbols constellation (number and placement of symbols) the more complicated the transmitter and receiver gets. This increases the development time of the device, component count and assembly cost.

Depending of the government rules and the standardization boards different trade-offs for phone systems will be the result. In Table 1.1 some of the most common phone systems are listed.

**Table 1.1: Some popular mobile radio systems (from [1])**

<table>
<thead>
<tr>
<th>Type</th>
<th>Year</th>
<th>Multiple Access</th>
<th>Frequency Band [MHz]</th>
<th>Modulation</th>
<th>Ch. BW [kHz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>NMT-900</td>
<td>1986</td>
<td>FDD</td>
<td>890-960</td>
<td>FM</td>
<td>12.5</td>
</tr>
<tr>
<td>AMPS</td>
<td>1983</td>
<td>FDD</td>
<td>824-894</td>
<td>FM</td>
<td>30</td>
</tr>
<tr>
<td>DECT</td>
<td>1993</td>
<td>TDD</td>
<td>1880-1900</td>
<td>GFSK</td>
<td>1728</td>
</tr>
<tr>
<td>GSM</td>
<td>1990</td>
<td>TDMA/FDMA/FDD</td>
<td>890-960</td>
<td>GMSK</td>
<td>200</td>
</tr>
<tr>
<td>DCS-1800</td>
<td>1993</td>
<td>TDMA/FDMA/FDD</td>
<td>1710-1880</td>
<td>GMSK</td>
<td>200</td>
</tr>
<tr>
<td>IS-54</td>
<td>1991</td>
<td>TDMA/FDMA/FDD</td>
<td>824-894</td>
<td>π/4 DQPSK</td>
<td>30</td>
</tr>
<tr>
<td>IS-95</td>
<td>1993</td>
<td>CDMA</td>
<td>824-894 1800-2000</td>
<td>QPSK/BPSK</td>
<td>1250</td>
</tr>
<tr>
<td>PDC</td>
<td>1993</td>
<td>TDMA/FDMA/FDD</td>
<td>810-1501</td>
<td>π/4 DQPSK</td>
<td>25</td>
</tr>
<tr>
<td>PHS</td>
<td>1993</td>
<td>TDMA/FDMA/TDD</td>
<td>1895-1907</td>
<td>π/4 DQPSK</td>
<td>300</td>
</tr>
<tr>
<td>WB-CDMA</td>
<td>-</td>
<td>CDMA/FDD</td>
<td>1920-2170</td>
<td>QPSK</td>
<td>5000</td>
</tr>
</tbody>
</table>

The first generation of radio systems, such as AMPS and NMT-900, was based on analog frequency modulation whereas the second and third generation are mainly based on digital modulation such as QPSK and GMSK. The access method is FDD (Frequency division duplex) for 1st generation whereas 2nd and 3rd generation uses TDD combined with TDMA (time-division multiple-access) and FDMA (frequency-division multiple-access). Also CDMA (code division multiple access) is used in 2nd generation. Wide band methods such as CDMA are more immune to fading but requires a strict power control to utilize the maximum users in the same band.
Analog FM and GMSK are constant envelope modulation methods which make the use of power efficient non-linear power amplifiers. QPSK have a varying envelope and requires a linear amplifier.

### 1.3.1 The modulation process

The purpose of the modulation is to transform the input signal (voice or data) into a signal suitable for transmission which means the signal should be band limited and suitable for detection at the receiver side. In modern transceivers all information including voice, data and signalling is digital. A block diagram of a modulator and a RF front-end is shown below. The modulator accommodates any linear modulation, and if an additional non-linear stage is inserted after the pulse-shaping it also accommodates non-linear modulation such as GMSK[1]

![Figure 1.7: The digital modulation process and the RF front end.](image)

The information from bit source is mapped into symbols. The numbers of symbols and their representation depends on the kind of modulation. For binary phase shift keying the alphabet of the symbols is “-1” and “+1” in such way that the sign of the carrier is determined by the bit information. For a 8-PSK signal the alphabet consist of 8 symbols. In this case three bits are used to map a symbol. The bandwidth of a signal can be reduced by increasing the alphabet size. The system however is more sensitive to noise as the alphabet increases.

After mapping the signal is interpolated and filtered. The interpolation can be perceived as a pulse amplitude modulation (PAM). The filtering determines the shape of the final spectrum. After the signal is modulated it is converted to an analog signal by the A/D converter and translated up in frequency by the mixer.

Interpolation is used to alleviate the requirements of the anti-aliasing filter but it is also necessary because the pulse-shaping sets the bandwidth of the modulated signal to a larger bandwidth than the symbol rate (the frequency of the symbols). This necessitates a higher sampling frequency.

The mapping process is often 2 dimensional in case as 8-PSK and QPSK. This means that each symbol consist of 2 values. Each value is then interpolated and filtered separately. The filtered signals are then up converted by two orthogonal local oscillator signals. This will be explained further in the following sections. Notice that the symbols are often described by complex numbers and that the interpolation and pulse shaping can be described fully by complex calculations [3].

### 1.3.2 IQ modulation

IQ modulation is widespread probably due to the history of radio and because it has a practical realisation. In the early radio day where amplitude modulation (AM) was used single side band modulation (SSB) was invented to save power. If only one side band is sent half the power can
be saved. This method is also more spectral efficient as only half the bandwidth is required. A block diagram of a SSB modulator is shown in Figure 1.8.

The input was typically a voice signal. It was fed to the two branches. In one branch the signal was up converted directly and in the other branch the signal was phase shifted 90 degrees before up-conversion. The RF local oscillator signals are 90 degrees out of phase. Both the LO signals and the signals, s and s’, to be up-converted are the Hilbert transformed of each other (see also Section 1.1.1 "Direct conversion" on Page 3). The up-converted signal from the two branches are added or subtracted to cancel either the upper or lower signal band.

Today all information, voice and other signalling is digital. This allows us to make two digital signals out of one by multiplexing the signal. So instead of calculating the quadrature signal as the 90 degree phase shifted signal of the input signal the two digital signals are used to modulate the two carriers in quadrature as showed in Figure 1.9. Side bands are not cancelled because the multiplexed signals are not related though the Hilbert transform but the bandwidth of the two new signals are half of the original signal. So the spectral efficiency would the same if the same amount of information has to be sent through a SSB transmitter (Figure 1.8).

In the following sections some examples of IQ modulation will be given. In the examples different types of filters and methods of plotting modulation signals will be presented.

Figure 1.8: Single side band modulator (SSB).

Figure 1.9: IQ-modulation.
1.3.3 QPSK modulation

QPSK modulation is characterized by its mapping scheme and the pulse-shaping. A constellation diagram of inphase and quadrature values are shown below in Figure 1.10. There are four symbols in the alphabet which can be characterized by a two bit number as shown in Table 1.2. For each symbol sent two bit are sent. The duration of a symbol period is twice the period of one bit ($T_s = 2T_b$) because there are used 2 bits per symbol. The carrier signal is described using the complex envelope of the corresponding symbol.

\[ s_{RF}(t) = \text{Re}[e^{j\cdot\omega_c t}] \cdot Cenv_n \quad \text{for } t \in \{n \cdot T_s < t \leq n + 1 \cdot T_s\} \]

\[ Cenv_n \in \{s_1, s_2, s_3, s_4\} \]

(1.25)

**Table 1.2: Symbol mapping and corresponding complex envelope.**

<table>
<thead>
<tr>
<th>Symbol</th>
<th>bits</th>
<th>Complex envelope (Cenv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1$</td>
<td>(0,0)</td>
<td>$\frac{1}{\sqrt{2}} \cdot e^{j \cdot \frac{\pi}{4}}$</td>
</tr>
<tr>
<td>$s_2$</td>
<td>(0,1)</td>
<td>$\frac{1}{\sqrt{2}} \cdot e^{j \cdot \frac{3\pi}{4}}$</td>
</tr>
<tr>
<td>$s_3$</td>
<td>(1,0)</td>
<td>$\frac{1}{\sqrt{2}} \cdot e^{-j \cdot \frac{\pi}{4}}$</td>
</tr>
<tr>
<td>$s_4$</td>
<td>(1,1)</td>
<td>$\frac{1}{\sqrt{2}} \cdot e^{-j \cdot \frac{3\pi}{4}}$</td>
</tr>
</tbody>
</table>

The phase of the carrier changes instantaneous for very symbol period, $T_s$. This result in a unlimited spectrum. To control the bandwidth a pulse shaping filter is used (see block diagram in...
Figure 1.9). Now the carrier have no discontinuities and the phase and amplitude will change during the symbol period.

\[ s_{RF}(t) = \text{Re}[e^{j\omega_c t} \cdot Cenv(t)] = \text{Re}[e^{j\omega_c t} \cdot (I(t) + jQ(t))] \]  \hspace{1cm} (1.26)

A common used filter is the raised cosine filter. This is a Nyquist filter [1] which has good properties concerning intersymbol interference (ISI). This can be seen at the impulse response in Figure 1.11. The black crosses are placed at a distance of one symbol period \( T_s \). If this pulse is sampled exactly at the symbol rate \( r_s = 1/T_s \) only one sample will be non zero. This is an advantage when several impulses are overlaid because they will not affect each other even though they are unlimited in time.

![Figure 1.11: Pulse response of a raised cosine filter (\( \alpha = 0.22 \)).](image)

![Figure 1.12: Frequency characteristic of the raised cosine filter.](image)
The raised cosine filter has two parameters, the symbol rate, $r_s$, and the bandwidth parameter, $\alpha$.

\[
H_{RC}(f) = \begin{cases} 
1 & 0 \leq |f| \leq \frac{1 - \alpha}{2T_s} \\
\frac{1}{2} \left[ 1 + \cos \left( \frac{\pi [2T_s |f| - 1 + \alpha]}{2\alpha} \right) \right] & \frac{1 - \alpha}{2T_s} < |f| \leq \frac{1 + \alpha}{2T_s} \\
0 & |f| > \frac{1 + \alpha}{2T_s}
\end{cases}
\] (1.27)

\[
h(t) = \left( \frac{\sin \frac{\pi t}{T_s}}{\pi t} \right) \frac{\cos \left( \frac{\pi \alpha t}{T_s} \right)}{1 - \left( \frac{4\alpha t}{2T_s} \right)^2}
\] (1.28)

Alphas determines the shape and bandwidth of the filter. Figure 1.12 shows the filter characteristic for three different alpha parameter ($\alpha = \{0, 0.22, 1\}$). The bandwidth of the filter is $r_s((1 + \alpha)/2)$ and the resulting modulated RF signal is the double $r_s(I + \alpha)$.

From the above it is seen that all information on the RF carrier signal is contained in the inphase and quadrature signal. For monitoring the received signal it is useful to plot the modulation signals in a coordinate system similar to the phase constellation diagram. Such a plot is shown in Figure 1.13. The corresponding inphase and quadrature signal are plotted against time in Figure 1.14.

![Figure 1.13: IQ constellation plot of QPSK modulated random signal with raised cosine filter ($\alpha=0.22$).](image-url)
In a transceiver system the raised cosine filtering is performed both on the transmitter and receiver side because it has some advantages concerning noise properties and linearity. In this case a root raised cosine filter is used on both the receiver and transmitter side. The filter is actually the square root of the raised cosine filter.

\[
H_{RRC}(f) = \sqrt{|H_{RC}(f)|} \cdot e^{j \frac{\arg(H_{RC}(f))}{2}}
\] (1.29)

Notice that the previous plots of the inphase and quadrature signal are based on a raised cosine filter and not a root raised cosine filter. This is done to show that the filtered signal actually passes through the constellation points. From Figure 1.13 it can be seen the trajectory some times passes though the origin. This means that the amplitude is zero and there is no carrier. The length of the vector from origin to a point on the trajectory in the IQ-trace diagram is the instant amplitude of the modulated RF carrier. The angle of the vector is the instant phase of the modulated signal. This can be seen from the equations (0.7) and (0.8) in Section 0.1.

1.3.4 Edge modulation

The EDGE modulation is an extension to the GSM system to provide high speed data when the transmission conditions allows it. Due to a closer position of the symbols a higher signal to noise ratio is needed to guarantee an unchanged bit error rate. The modulation is a phase shift keying (PSK) method with subsequent symbol rotation to avoid passing through origin in the IQ-tra-
jectory plot and thereby avoiding zero amplitude. The constellation diagram is shown in Figure 1.15. Which corresponds to the following symbol mapping table.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Symbol parameter, k</th>
</tr>
</thead>
<tbody>
<tr>
<td>s₀</td>
<td>(1,1,1)</td>
</tr>
<tr>
<td>s₁</td>
<td>(0,1,1)</td>
</tr>
<tr>
<td>s₂</td>
<td>(0,1,0)</td>
</tr>
<tr>
<td>s₃</td>
<td>(0,0,0)</td>
</tr>
<tr>
<td>s₄</td>
<td>(0,0,1)</td>
</tr>
<tr>
<td>s₅</td>
<td>(1,0,1)</td>
</tr>
<tr>
<td>s₆</td>
<td>(1,0,0)</td>
</tr>
<tr>
<td>s₇</td>
<td>(1,1,0)</td>
</tr>
</tbody>
</table>

The mapping is described based on the symbol parameter in Table 1.3.

\[
s_k = e^{j \frac{2\pi \cdot k}{8}}
\]

(1.30)

The symbol is rotated in steps of \(3\pi/8\).

\[
\tilde{s}_k = s_k \cdot e^{j \frac{3\pi}{8} \cdot k}
\]

(1.31)

The puls-shape filter is a linearised Gauss pulse (see definition in Appendix B). The impulse response can be seen in Figure 1.16. For the comparison the Gauss pulse used in GMSK is plotted too. The spectrum of the EDGE signal is designed to use the same bandwidth as the GMSK signal for compatibility. The reason that the two signal are not equal is that one filter is used for beings.

![Figure 1.16: Gauss pulse and linearised Gauss pulse.](image-url)
a linear modulation (EDGE, 8PSK) and the other one is used for a non-linear modulation[1] (GMSK). The Gauss pulse (BT=0.21) is used to filter the phase of the carrier signal whereas the linearised Gauss pulse is used filter the inphase and quadrature signal.

To avoid any doubts concerning the definition of the resulting modulated signal in the radio standard (EDGE/GSM) the definition is formulated based on a filtered impulse response. A mapped and rotated symbol is multiplied by a delta Dirac pulse and filtered by an ideal time continuous filter. The result is a complex envelope that can be used for an IQ up-conversion (see Figure 1.17 and the radio specification[4]). It is up to the designer to compensate for sample-

![Diagram](image-url)

**Figure 1.17: Principle of EDGE modulation definition.**

hold effects caused by D/A converters and finite length FIR filters used in the base band processor. The quality of the modulated signal of a practical transceiver have to fulfil some requirements compared to the ideal defined signal. These requirements are based on the Error Vector Magnitude (EVM) and spectral leakage requirements. These measures are explained in the next chapter (Chapter 1.3.5).

### 1.3.5 Measuring the signal quality

The modulated signal is exposed to non-linearities and noise propagating from the transmitter to the receiver. To determine the signal quality some quality measures are needed. The most important measures are the Error Vector Magnitude (EVM), the Adjacent Channel Power (ACP) and the Bit Error Rate (BER). The latter is a statistic of the number of bits received correctly. A typical number for data communication is 10E-6 which means that 1bit out of a million is detected false. This is of course not acceptable for data communication but then a protocol with checksum, re-transmission and error correction codes[3] can be used. The principle of measuring the bit error rate is shown in Figure 1.18. A random bit sequence is generated and sent though the modulator and the channel. At the receiver side the signal is demodulated and the bits detected. The received bits are compared to the bits sent and a error statistic is calculated. The BER is a important figure for evaluation the overall performance of a system.

The Error Vector Magnitude (EVM) is a number for the quality of the modulated signal and can be used to evaluated the linearity of the transmitter. The calculation can be carried out as suggested in Figure 1.19. A long modulated signal is generated and sent through the non-linearity (or channel). The non-linearity could for instance be the power amplifier in the transmitter. The degraded signal is then filtered by a filter corresponding to the one used in the receiver. This signal is aligned in time, phase and amplitude relative to a undistorted signal. This necessary because the channel (or non-linearity) introduce a delay and attenuation of the signal besides distortion and noise. For the phase alignment the complex correlation between the complex envelopes signals can be calculated. The phase of the coefficient specifies the phase difference between the signals. The amplitude alignment is performed by minimizing the RMS value of the difference between the complex envelope signals. The difference signal is called the error vec-
tor signal. Minimizing the RMS value of this signal is actually the same as minimizing the output of the system in Figure 1.19 which is the error vector magnitude. The error vector magnitude is the RMS value of the error vector signal divided by the RMS value of the undistorted signal. This value specified in percent. A typically maximum value in a radio system is 10%.

The above explained EVM measure defines the inband quality. The measure does not specify spectral components appearing outside the radio channel liable to disturb communication in the neighbour channels.

The adjacent channel power (ACP) measures the power in the neighbour relative to the power in the wanted channel. The principle of the calculation is shown in Figure 1.20. Depending on

---

**Figure 1.18:** Measurement of Bit Error Rate (BER).

**Figure 1.19:** Calculation of the error vector and the error vector magnitude (EVM).

**Figure 1.20:** Calculation of adjacent channel power.
ification a 30kHz filter is used. Also the frequency offset of the filter can be varied and several requirements for different offset must be respected. Notice that all the suggested figures of signal quality can be simulated efficiently using complex envelope (see “Complex envelope representation” on Page 8).
Chapter 2

Modelling and Characterising Non-Linearities

In this chapter some models and measures for non-linearities are presented. The focus will be on models based on power series and limiters as they are suitable to model functional blocks in a transceiver such as amplifiers, mixers and limiters. Power series and limiters are memory less models which are unable to distort the phase of the input signal. The bandpass amplifier model is similar to the memoryless non-linearity but it has phase non-linearity that dependent on the instantaneous input amplitude as we will see in the following sections.

Another important type of non-linearity is the non-linearities with memory. An example is the smith-trigger where the comparator level depends on a state variable. A generalised type of smith-trigger is presented at the end of this chapter. Functions describing the bandpass behaviour in cartesian coordinates are given.

2.1 Harmonic distortion

Consider the non-linear system e.g. an amplifier, in Figure 2.1 with sine signal as input.

\[ x(t) = A \cos(\omega_c t + \Theta) \]  

(2.1)

\[ y(t) = b_0 + \sum_{n=1}^{N} b_n \cos(n\omega_c t + \Theta_n) \]  

(2.2)

Figure 2.1: Non-linear system.

The output of the system will in general be a weighted sum of sinusodials,
where the amplitude and the phase depend on the input signal and the non-linearities. The small signal gain is defined as the ratio of the output and the input amplitude at the fundamental frequency at a very low input amplitudes.

\[ G_{\text{lin}} = \left. \frac{b_1(A)}{A} \right|_{A \to 0} \]  

As the input voltage increases the gain will change. A gain increase is denoted as gain expansion, whereas a reduction is denoted as gain compression. Amplifiers will normally compress due to saturation effects for higher drive levels. For power amplifiers the 1-dB compression point is normally specified. This measure indicates the input level for the actual gain is 1-dB lower than the small signal gain.

\[ A_{1dB} \left( \frac{b_1(A_{1dB})}{G_{\text{lin}} \cdot A_{1dB}} \right) = 10^{\frac{-1}{20}} \]  

The compression is normally specified in dBm (decibel relative to 1mW) assuming that the voltage is delivered to a 50Ω load.

Functions that describes the output amplitude and the phase behaviour of the fundamental as a function of the input amplitude are designated “amplitude to amplitude conversion function” (AM/AM conversion) and “amplitude to phase conversion function” (AM/PM conversion).

Harmonic distortion is characterised as the ratio between the amplitude of a harmonic spectral component and the amplitude of the fundamental frequency.

\[ HD_i = \left| \frac{b_i(A)}{b_1(A)} \right| \]  

Harmonic distortion intercept points are defined for small signals as higher order non-linearities are neglected. The n-order intercept point is defined as the intercept point of the extrapolated power of the nth harmonic and the power of fundamental spectral component when power is measured in Decibel.

\[ IP_{nh} = \left. \frac{P_{dB,n}(A_0) - P_{dB,1}(A_0)}{\left( \frac{dP_{dB,1}}{dA} - \frac{dP_{dB,n}}{dA} \right)_{A \to 0} + P_{in}} \right|_{A \to 0} \]  

where \( P_{in} \) is the power of the input signal \( 10 \cdot \log(A_0^2/2) \) in dB and the index \( h \) denotes a intercept point for harmonic distortion (and not intermodulation). \( P_{dB,1}, P_{dB,n} \) are the power of the fundamental frequency component and the nth harmonic \( P_{dB,i} = 10 \cdot \log(b_i^2/2) \).

The total harmonic distortion is the ratio of the power of all higher order harmonics and the power of the fundamental.

\[ THD = \sum_{i=2}^{\infty} \frac{b_i^2}{b_1^2} \leq 2 \]  

(2.7)
2.2 Intermodulation

Intermodulation measures are defined only for low input amplitudes where higher order effects do not influence the amplitude of the spectral components. To characterise intermodulation a two tune input with equal amplitude is used.

\[ x(t) = A[\cos(\omega_1 t) + \cos(\omega_2 t)] \]  
\[ (2.8) \]

Spectral components in general arise at all combinations of the two input frequencies.

\[ \omega = n\omega_1 + m\omega_2 \quad n, m \in Z_0 \]  
\[ (2.9) \]

The amplitude of the spectral component at the frequency \( n\omega_1 + m\omega_2 \) is denoted \( v_{n,m} \). Using this notation the 2nd and 3rd order intermodulation are defined as the amplitude of the actual mixed spectral component and the input amplitude.

\[ IM_{|n|+|m|} = \left| \frac{v_{m,n}}{v_{1,0}} \right| \quad \omega = n\omega_1 + m\omega_2 \]  
\[ (2.10) \]

\[ IM_2 = \left| \frac{v_{1,1}}{v_{1,0}} \right| \]  
\[ (2.11) \]

\[ IM_3 = \left| \frac{v_{2,-1}}{v_{1,0}} \right| \]  
\[ (2.12) \]

The order of the intermodulation product is defined as the sum of the frequency indexes.

\[ \text{order}_{IM} = |n| + |m| \]  
\[ (2.13) \]

Intermodulation intercept points are defined as the point where the fundamental and the intermodulation product are equal in amplitude, assuming low input amplitude and neglecting higher order effects.

\[ IP_{|n|+|m|} = A\left\{ \left| \frac{v_{1,0}}{v_{1,0}} \right| = \left| \frac{v_{m,n}}{v_{1,0}} \right| \right\} \quad A \text{ is small} \]  
\[ (2.14) \]

\[ IP_2 = A\left\{ \left| \frac{v_{1,0}}{v_{1,0}} \right| = \left| \frac{v_{1,1}}{v_{1,0}} \right| \right\} \]  
\[ (2.15) \]

\[ IP_3 = A\left\{ \left| \frac{v_{1,0}}{v_{1,0}} \right| = \left| \frac{v_{2,1}}{v_{1,0}} \right| \right\} \]  
\[ (2.16) \]

As the input level increases, higher order terms will change the actual gain. In case of an amplifier the gain will normally decrease. The point, where gain is 1dB lower than the small signal gain is called the 1-dB compression point.

\[ A_{1dB} \quad \text{where} \quad \left( \frac{b_1(A_{1dB})}{A_{1dB}} \right) / G_{lin} = a \log \frac{1}{20} \]  
\[ (2.17) \]
2.3 Polynomial non-linearities

Polynomial non-linearities with no memory can be described by a n-order polynomial, the instant output value as a function of the instant value of the input.

\[ y = f(x) = a_0 + \sum_{i=1}^{N} a_i x^i \]  

(2.18)

Many amplifiers can be approximated with a polynomial non-linearity as long as the frequency is low enough. Furthermore the polynomial non-linearities provides a good understanding of distortion measures because explicit formulas for the distortion figures are obtained. To observe the response to a single input signal at the fundamental frequency a signal, \( A \cos(\omega_c t + \Theta) \), is inserted and some trigonometric relations are used to form a sum of cosines at the frequencies 1\( \omega \), 2\( \omega \) to \( N\omega \). Equation (2.18) can be written as shown below.

\[ y = f(A \cos(\omega_c t + \Theta)) = b_0(A) + \sum_{n=1}^{N} b_n(A) \cos(n\omega_c t + n\Theta) \]  

(2.19)

Each coefficient \( b_i(A) \) in equation (2.19) specifies the amplitude of the \( i \)th harmonic spectral component and is a polynomial function of the amplitude of the input signal and the coefficients \( a_i \). The coefficients can either be found solving the integrals in equation (2.41) and (2.42) as shown in Chapter 2.4 or using trigonometrical manipulations in equation (2.19). Depending of the what kind of non-linearity we want to characterize different figures and coefficient will be relevant. If the non-linearity represents a power amplifier \( b_1 \) determines parameters such as gain and compression point of the amplifier.
For a 9rd order polynomial \((N = 9)\) equation (2.19) evaluates to

\[
y = a_0 + \frac{1}{2}a_2 A^2 + \frac{3}{8}a_4 A^4 + \frac{5}{16}a_6 A^6 + \frac{35}{128}a_8 A^8
\]

\[
+ \left( a_1 A + \frac{3}{4}a_3 A^3 + \frac{5}{8}a_5 A^5 + \frac{35}{64}a_7 A^7 + \frac{63}{128}a_9 A^9 \right) \cos (\omega t + \Theta)
\]

\[
+ \left( \frac{1}{2}a_2 A^2 + \frac{1}{2}a_4 A^4 + \frac{15}{32}a_6 A^6 + \frac{7}{16}a_8 A^8 \right) \cos (2\omega t + 2\Theta)
\]

\[
+ \left( \frac{1}{4}a_3 A^3 + \frac{5}{16}a_5 A^5 + \frac{21}{64}a_7 A^7 + \frac{21}{64}a_9 A^9 \right) \cos (3\omega t + 3\Theta)
\]

\[
+ \left( \frac{1}{8}a_4 A^4 + \frac{3}{16}a_6 A^6 + \frac{7}{32}a_8 A^8 \right) \cos (4\omega t + 4\Theta)
\]

\[
+ \left( \frac{1}{16}a_5 A^5 + \frac{7}{64}a_7 A^7 + \frac{9}{128}a_9 A^9 \right) \cos (5\omega t + 5\Theta)
\]

\[
+ \left( \frac{1}{32}a_6 A^6 + \frac{1}{16}a_8 A^8 \right) \cos (6\omega t + 6\Theta)
\]

\[
+ \left( \frac{1}{64}a_7 A^7 + \frac{9}{128}a_9 A^9 \right) \cos (7\omega t + 7\Theta)
\]

\[
+ \frac{1}{128}a_8 A^8 \cos (8\omega t + 8\Theta) + \frac{1}{256}a_9 A^9 \cos (9\omega t + 9\Theta)
\]

From this equation some general properties characteristic for non-linearities are observed.

1) The phase and the frequency of the \(n\)th harmonic is multiplied by the order of the harmonic.

2) At low input amplitudes high order terms for the harmonics can be neglected. In this case the amplitude (or power) of the \(n\)th harmonic increases by \(n \cdot \text{dB}\) as the input increases by \(\text{dB}\).

Many of the distortion measures are defined for low input amplitudes. In this case higher order products can be neglected and equation (2.20) reduces to.

\[
y = a_0 + \frac{1}{2}a_2 A^2 + a_4 A \cos (\omega t + \Theta) + \frac{1}{2}a_2 A^2 \cos (2\omega t + 2\Theta)
\]

\[
+ \frac{1}{4}a_3 A^3 \cos (3\omega t + 3\Theta) + \frac{1}{8}a_4 A^4 \cos (4\omega t + 4\Theta)
\]

\[
+ \frac{1}{16}a_5 A^5 \cos (5\omega t + 5\Theta) + \frac{1}{32}a_6 A^6 \cos (6\omega t + 6\Theta)
\]

\[
(2.21)
\]

\[
+ \frac{1}{64}a_7 A^7 \cos (7\omega t + 7\Theta) + \frac{1}{128}a_8 A^8 \cos (8\omega t + 8\Theta)
\]

\[
+ \frac{1}{256}a_9 A^9 \cos (9\omega t + 9\Theta)
\]

**2.3.1 Harmonic distortion**

Now the different distortion measures can be expressed explicit.
SECTION 2.3 - POLYNOMIAL NON-LINEARITIES

Harmonic distortion:

\[ HD_{2h} = \frac{b_2}{b_1} \bigg|_{A_{\text{small}}} = \frac{1}{2} A \frac{a_2}{a_3} \]  \hspace{1cm} (2.22)

\[ HD_{3h} = \frac{b_3}{b_1} \bigg|_{A_{\text{small}}} = \frac{1}{4} A \frac{a_3}{a_3} \]  \hspace{1cm} (2.23)

Intercept points:

\( (b_1 = b_2) \Rightarrow \left( a_1 IP_{2h} = \frac{1}{2} a_2 (IP_{2h})^2 \right) \Rightarrow \)  \hspace{1cm} (2.24)

\[ IP_{2h} = \frac{a_1}{a_2} \]  \hspace{1cm} (2.25)

\( (b_1 = b_3) \Rightarrow IP_{3h} = 2 \sqrt[3]{\frac{a_1}{a_3}} \)  \hspace{1cm} (2.26)

The small signal gain is defined as the ratio of the input and output amplitude at a very low input signal level.

\[ G_{\text{lin}} = \left. \frac{b_1(A)}{A} \right|_{A \to 0} = a_1 \]  \hspace{1cm} (2.27)

If a 3rd order non-linearity is assumed, the n-dB compression point can be written as

\[ A_{\text{ndB}} = \sqrt[\frac{n}{20}]\left[ \frac{4}{3} \cdot \frac{a_1}{a_3} \cdot \frac{1 - u}{u} \right] \]  \hspace{1cm} u = 10^{\frac{n}{20}} \hspace{1cm} (2.28)

2.3.2 Intermodulation

The intermodulation is calculated using two tunes as input. See “Intermodulation” on Page 23.

\[ x(t) = A \left[ \cos(\omega_1 t) + \cos(\omega_2 t) \right] \]  \hspace{1cm} (2.29)
Then equation (2.19) evaluates to

\[ y = f(A[\cos(\omega_1 t) + \cos(\omega_2 t)]) \]
\[ = a_0 + a_2 A^2 \]
\[ + \left( 1a_1 A + \frac{9}{4}a_3 A^3 \right)[\cos(\omega_1 t) + \cos(\omega_2 t)] \]
\[ + \frac{1}{2} a_2 A^2[\cos(2\omega_1 t) + \cos(2\omega_2 t)] \]
\[ + \frac{1}{4} a_3 A^3[\cos(3\omega_1 t) + \cos(3\omega_2 t)] \]
\[ + 1a_2 A^2[\cos(\omega_1 t + \omega_2 t) + \cos(\omega_1 t - \omega_2 t)] \]
\[ + \frac{3}{4} a_3 A^3[\cos(2\omega_1 t + \omega_2 t) + \cos(2\omega_1 t + \omega_2 t)] \]
\[ + \frac{3}{4} a_3 A^3[\cos(2\omega_1 t - \omega_2 t) + \cos(\omega_1 t - 2\omega_2 t)] \]  

Equation (2.30)

Now the intermodulation can be expressed explicit

\[ IM_2 = \left| \frac{v_{1,1}}{v_{1,0}} \right| = \frac{a_2 A^2}{|a_1 A|} = \frac{a_2}{a_1} A \]  

(2.31)

\[ IM_3 = \left| \frac{v_{2,-1}}{v_{1,0}} \right| = \frac{\frac{3}{4} a_3 A^3}{|a_1 A|} = \frac{3}{4} \frac{a_3}{a_1} A \]  

(2.32)

The 2nd and 3rd order intermodulation intercept points are calculated as

\[ [a_1 IP_{2i} = a_2 (IP_{2i})^2] \Rightarrow IP_{2i} = \frac{a_1}{a_2} \]  

(2.33)

\[ [a_1 IP_{3i} = \frac{3}{4} a_3 (IP_{3i})^3] \Rightarrow IP_{3i} = 2 \sqrt[3]{\frac{a_1}{3a_3}} \]  

(2.34)

\( IP_{3i} \) is the necessary amplitude of the two sinusodials signals at the input of the non-linearity to make the 1st and the 3rd order intermodulation components equal. Comparing equation (2.25) and (2.33) one see that they are proportional.

\[ IP_{2h} = 2 \cdot IP_{2i} \]  

(2.35)

The same is the case for equation (2.26) and (2.34).

\[ IP_{3h} = \sqrt[3]{3} \cdot IP_{3i} \]  

(2.36)
2.4 Memory-less bandpass non-linearities

A memory-less non-linearity can be described as an instantaneous function of the input signal. The result of a low frequency baseband signal can be described directly using the function of the non-linearity.

\[ y(t) = F[x(t)] \quad (2.37) \]

This, however, is computation insufficient for bandpass signals because every period of the carrier has to be evaluated even though the phase and the amplitude varies slowly. In this case it more efficient to use a description based on complex envelope signals.

First, assume that the phase and amplitude of the carrier are constant.

\[ x(t) = A \cdot \cos(\omega_c t + \Theta) \quad (2.38) \]

where

- \( A \) is the amplitude
- \( \Theta \) is the phase

Then the output signal of the non-linearity will be a periodic signal which can be described by a Fourier expansion. To simplify the equations the argument of the cosine is substituted

\[ y(\alpha) = F[A \cdot \cos(\alpha)] \quad (2.39) \]

where

\[ \alpha = \omega_c t + \Theta \]

The output signal is then described by

\[ y(\alpha) = a_0 + \sum_{k=1}^{\infty} (a_k \cos(k\alpha) + b_k \sin(k\alpha)) \quad (2.40) \]

where

\[ a_k = \frac{1}{\pi} \int_{0}^{2\pi} F[A \cdot \cos(\alpha)] \cos(k\alpha) d\alpha \quad (2.41) \]

and

\[ b_k = \frac{1}{\pi} \int_{0}^{2\pi} F[A \cdot \cos(\alpha)] \sin(k\alpha) d\alpha \quad (2.42) \]

Equation (2.40) consist of a DC term plus a sum of harmonic components at the frequencies \( \omega_c, 2\omega_c, ..., n\omega_c \) where \( k \) is the index of the harmonic. If \( A \) and \( \Theta \) are varying slowly compared to the carrier period then the harmonics will change into frequency bands placed around the harmonics of the carrier frequency. This is also seen due to the fact that the coefficients in equation (2.41) and (2.42) become time varying. In most cases the output signal at the fundamental frequency is of interest. In this case equation (2.40) can be abbreviated to

\[ y(t) \equiv a_1 \cos(\omega_c t + \Theta) + b_1 \sin(\omega_c t + \Theta) \quad (2.43) \]

where \( a_1 \) and \( b_1 \) are functions of \( A \) and \( \Theta \).
According to [33] any ordinary function, $F$, will result in $b_k$ being equal to zero. This means that there is no phase distortion (AM/PM). This is the case for polynomial non-linearities and limiters. Functions with hysteresis are able to generate a non-zero coefficient $b_k$ which means that phase distortion is generated[33]. Equation (2.43) is also used to model the distortion including the phase distortion that is experienced in bandpass amplifiers. It should also be noted that in the case of a narrow bandpass amplifier $a_k$ and $b_k$ are only amplitude dependent even though the input signal is varying both in amplitude and phase. This is equivalent to a non-frequency dependent non-linearity. If a wide band amplifier is modelled $a_k$ and $b_k$ are also phase dependent.

The complex function $(1/A)[a_1(A) + jb_1(A)]$ is in the control literature called the describing function. Equation (2.41) is the Chebyshev transform of $F(x)$. In this text the output of the non-linearity is often described by the complex envelope which according to Section 0.1.2 can be written as follows.

$$[a_1(A) - jb_1(A)]e^{\Theta}$$

(2.44)

### 2.4.1 Complex envelope bandpass non-linear simulation model

The non-linearities described by equation (2.43), can efficiently be simulated using complex envelope simulation as shown in Figure 2.2. The model is based on polar coordinates. Therefore if the cartesian representation of the non-linearity in equation (2.43) is used it must be converted to polar form.

$$f(A) = |a_1(A) + jb_1(A)|$$

(2.45)

$$g(A) = \text{arg} [a_1(A) + jb_1(A)]$$

(2.46)

Figure 2.2 shows that the phase and the amplitude is extracted of the input signal. The amplitude is then used to calculate the new amplitude using the AM/AM distortion function. The phase distortion is calculated using the AM/PM function. The AM/AM and AM/PM can be functions based on the equations (2.41) and (2.42) or they can be table based origination from measurements of a physical device.
2.4.2 The limiter non-linearity

The general limiter function is defined as

\[
\frac{L_{out} \cdot \text{sgn}(x)}{\left[1 + \left(\frac{L_{in}}{|x|}\right)^s\right]^\frac{1}{s}}
\]

(2.47)

where

- \(s\) is the knee sharpness
- \(L_{in}\) is the input limit.
- \(L_{out}\) is the maximum output amplitude.

The function is depicted in Figure 2.3.

![Figure 2.3: Limiter function (s=10, Lin=Lout=1).](image)

If a bandpass signal is applied to the limiter the first zonal output (the bandpass signal at the fundamental frequency) can be described by equation (2.43). In the case of a hard limiter (\(L_{in}=0\), \(s\) is not defined) the output is given by the following equations

\[
\frac{1}{A} a_1(A) = \frac{2L_{out}}{\pi L_{in}} \left[\text{asinc}\left(\frac{L_{in}}{A}\right) + \frac{L_{in}}{A} \sqrt{1 - \left(\frac{L_{in}}{A}\right)^2}\right] \quad A \geq L_{in} \quad (2.48)
\]

\[
\frac{1}{A} a_1(A) = \frac{L_{in}}{A} \quad A < L_{in} \quad (2.49)
\]

\[
b_1 = 0 \quad (2.50)
\]

\(b_k\) is equal to zero which means that no phase distortion is present.
2.4.3 A non-linearity with hysteresis

Now consider the non-linearity in Figure 2.4. This is a non-linearity with hysteresis. The first zonal output is given by

\[
\frac{1}{A} a_1(A) = \frac{k}{\pi} \left[ \frac{\pi}{2} + \arcsin \left( 1 - \frac{2D}{A} \right) + 2 \left( 1 - \frac{2D}{A} \right) \sqrt{\frac{D}{A} \left( 1 - \frac{D}{A} \right)} \right]
\]  
(2.51)

\[
\frac{b_1}{A} = \frac{(-4)kD}{\pi A} \left( 1 - \frac{D}{A} \right)
\]  
(2.52)

Here it is seen that \(b_1\) is non zero which corresponds to a phase distortion of the output signal. Depending on the slope \(k\) the non-linearity exhibits more or less phase distortion.

![Figure 2.4: A non-linearity with hysteresis.](image)
Chapter 3

The Power Amplifier

The purpose of an RF power amplifier is to amplify the input signal and provide the necessary power to a load with a minimum loss of power in the amplifier itself. The output power and the gain will be important specifications. But also the power efficiency, the sensitivity to load variations and stability are important properties.

In this chapter the theory for tuned amplifiers will be presented. Measures such as output power and efficiency will be calculated as a function of the conduction angle. In this way the PA operation class (A,B,C) can be treated at the same time. First the class A and class B amplifier is presented separately to ease the introduction. An example of an amplifier is shown below (Figure 3.1). The amplifier consists of a transistor, two inductors, three capacitors and a bias circuit. The bias circuit sets the DC operation point of the transistor. The voltage depends on the class of operation, that is the waveform of the current in the drain.

In class A operation, the transistor is always on whereas in class B, the transistor is only on half of the period of the sinusoidal. $C_2$ and $C_3$ are used as DC isolation and should have a low impedance at the carrier frequency compared to other impedances in the amplifier. The inductors $L_2$ and $L_3$ are AC isolators and provides DC voltages and currents. The resonator $L_1$, $C_1$ eliminates the harmonics of the drain current and ensures that the voltage over the load is a sinusoidal. The shape of the drain current is non-sinusodial except for the case of class A operation. Without the resonance circuits the voltage across the load would adopt the shape of the drain current. The class of operation is determined by the level and the shape of the drain cur-

![Figure 3.1: Power amplifier stage.](image)
This also determine the theoretically achievable power efficiency of the amplifier. In the next chapter the different classes of operation will be described.

### 3.1 Classes of operation

#### 3.1.1 Class A

Class A operation is the most linear class of operation because the transistor is active during the hole period of the carrier. A schematic of a class A amplifier is shown in Figure 3.2. It is similar to the one in Figure 3.1 except for the LC-tank. No resonator \((L_1,C_1)\) is necessary to support class A operation since the drain current is ideally sinusodial as depicted in Figure 3.3.

![Figure 3.2: Power amplifier stage for class A operation.](image)

The quiscent point of the amplifier is determined by the gate voltage and the power supply. The drain quiescent voltage is equal to the power supply as the DC voltage drop across \(L_2\) is zero. The drain quiescent current is determined by gate voltage and must be greater than maximum amplitude of the wanted AC drain current to ensure that the transistor is always active. The gate voltage is set by the bias circuit.

The relationship between the changes in drain voltage and current is determined by the load as the AC current flows solely though the load. The inductor \(L_2\) is a AC blocking device with an infinite AC impedance. The drain current and voltage can be written as:

\[
I_d = I_{dQ} + I_p \sin(\omega t) \quad \quad I_d \geq I_{d, sat}
\]

\[
V_d = V_{dd} - R_L \cdot I_p \sin(\omega t) \quad \quad V_{max} \geq V_d \geq V_{vsat}
\]  

(3.1)
The drain voltage limited downwards by the drain-bulk diode and upwards by the breakdown voltage. The working condition of the transistor is depicted in Figure 3.4.

The straight line is the loadline which can be derived from equation (3.1)

\[ I_d = -\frac{1}{R_L}V_d + \left( I_{dQ} + \frac{V_{dd}}{R_L} \right) \]  

(3.2)

The drain quiescent current can be derived by inserting a known point \((V_{max},0)\) on the load line in the above equation.

\[ I_{d}(V_{max}) = -\frac{1}{R_L}V_{max} + \left( I_{dQ} + \frac{V_{dd}}{R_L} \right) = 0 \Rightarrow \]

\[ I_Q = \frac{V_{max} - V_{dd}}{R_L} \]  

(3.3)

To avoid clipping the following conditions must be fulfilled.

\[ I_p R_L \leq v_{out \max} = \min(V_{dd} - V_{sat}, V_{max} - V_{dd}) \]  

(3.4)

The output power is described by the peak current in the load.

\[ P_{out} = \frac{I_p^2}{2} \cdot R_L \]  

(3.5)

The (drain) efficiency of the amplifier is defined as the ratio of the output power and the total dissipated power.

\[ \eta_{\text{drain}}(I_p) = \frac{P_{out}}{P_{DC}} = \frac{\frac{I_p^2}{2} \cdot R_L}{(V_{dd} \cdot I_Q)} = \frac{I_p^2 R_L}{2 V_{dd} I_Q} \]  

(3.6)
Given the maximum voltage swing over the load is achieved when $V_{\text{out,max}}$ equation (3.4) is maximised, which implies that $V_{dd}$ is the average of $V_{sat}$ and $V_{\text{max}}$.

$$V_{dd} = \frac{V_{\text{max}} + V_{sat}}{2} \quad (3.7)$$

Using this in equation (3.3) and (3.4) the efficiency is calculated to

$$\eta_{\text{drain, max}} = \frac{1}{2} \frac{(V_{\text{max}} - V_{sat})^2}{V_{\text{max}}^2 - V_{sat}^2} \leq \frac{1}{2} \quad (3.8)$$

### 3.1.2 Class B

In class B operation the drain current of the transistor has the shape of a rectified sinusoidal as shown in Figure 3.6. The employed amplifier is the same as the one in Figure 3.1 and is repeated in Figure 3.5.

![Figure 3.5: Class B amplifier.](image)

![Figure 3.6: Drain current of a transistor in class B operation.](image)

A parallel resonator set at the fundamental frequency in parallel to the load is necessary to ensure a sinusoidal voltage at the output. It will short all harmonic of the drain current and prevent them from going into the load. Only the fundamental of the drain current enters the load.
The fundamental of the drain current and the voltage of the load will be in phase because the resonator has an infinite impedance at the fundamental frequency and the load is resistive.

\[
P_{\text{load}} = \frac{1}{2\pi} \int_{0}^{\pi} R_L \cdot (I_p \sin(\varphi))^2 d\varphi = \frac{I_p^2 R_L}{4}
\]  

(3.9)

The average current though the amplifier stage is the average of the drain current as no DC current flows through the \( C_2 \). The power consumed by the amplifier is

\[
P_{\text{DC}} = V_{dd} \cdot \frac{1}{2\pi} \int_{0}^{\pi} I_p \sin(\varphi) d\varphi = \frac{V_{dd} I_p}{\pi}
\]  

(3.10)

And the efficiency is therefore

\[
\eta_{\text{drain}} = \frac{P_{\text{load}}}{P_{\text{DC}}} = \frac{I_p R_L}{V_{dd}} \cdot \frac{\pi}{4} \leq \frac{\pi}{4} = 0.785
\]  

(3.11)

The efficiency is below 78.5% because \( I_p \) is limited by the clipping condition.

\[
I_p R_L \leq V_{dd}.
\]  

(3.12)

### 3.1.3 Classes described by the conduction angle

A common way to describe the operation of the PA is to use the conduction angle defined as half the angle of the angle that the transistor is active. This will be explained further below. Class C differs from class B in the sense that conduction angle can is less than \( \pi/2 \) opposed to class B where the angle is always \( \pi/2 \). In the following we will derive the equations for all classes of operation. The parameter that characterize the class of operation is the conduction angle, \( \Theta_c \). The drain current can be seen as a rectified sinusodial with an offset (see Figure 3.7). The drain current is described by the following equation.

\[
I_d(\Theta, \Theta_c) = \begin{cases} 
I_{dd}(\Theta_c) \sin(\Theta) - I_{dQ}(\Theta_c) & \text{If positive} \\
0 & \text{Otherwise} 
\end{cases}
\]  

(3.13)

In Figure 3.7 the drain current is showed for different conduction angles. The drain current of a class A amplifier is a intact sinusodial, but for all other modes of operation the current is clipped in the bottom. Also notice that the peak current for small conduction angles are high. This is due to the fact that if a certain power has to be delivered to the load in a short time then the peak current has to be higher. The peak current will actually converge towards infinity as the conduction angle approaches zero as we will see later.

The conduction angle is defined as the half angle interval where the transistor is active. This way the conduction angle, \( \Theta_c \), can be expressed as a function of the parameters in equation (3.13).

\[
I_d\left(\frac{\pi}{2} - \Theta_c\right) = 0
\]  

(3.14)
One can also use the above equation to describe the quiescent current, $I_{dQ}$, in terms of the conduction angle and the drive current.

$$I_{dQ} = \cos(\Theta_c) \cdot I_{dd}$$  \hfill (3.15)

The same argument as for class B is applies to class C concerning the loading of the harmonics of the drain current. Only the fundamental frequency of the current is lead though the load and causes a voltage swing across the drain and the load. The other harmonics are shorted by the resonator. Below the fundamental of the drain current and the amplitude of sinusodial on the drain and the load is calculated.

$$I_{ac} = \frac{1}{\pi} \cdot \frac{2\pi}{\pi/2 + \Theta}$$

$$= \frac{1}{\pi} \cdot \int_{\pi/2 - \Theta}^{\pi/2 + \Theta} (I_{dd} \sin(y) - I_{dQ}(\Theta_c)) \cdot \sin(y) dy$$  \hfill (3.16)

$$= I_{dd} \cdot \frac{2\Theta_c - \sin(2\Theta_c)}{2\pi}$$

$$= I_{dd} \cdot \frac{\Theta_c - \cos(\Theta_c)\sin(\Theta_c)}{\pi}$$

The maximum voltage swing over the load is not allowed to go beyond the supply voltage, $V_{dd}$, to avoid clipping.

$$v_{load} = R_L \cdot I_{ac} \leq V_{dd}$$  \hfill (3.17)
This leads to a maximum current in the load, $I_{ac,max}$ and a maximum for $I_{dd}$ called $I_{dd,max}$:

$$v_{load,max} = I_{ac,max} \cdot R_L = I_{dd,max} \cdot \frac{2\Theta_c - \sin(2\Theta_c)}{2\pi} \cdot R_L = V_{dd}$$ (3.18)

$$I_{ac,max} = \frac{V_{dd}}{R_L}$$ (3.19)

$$I_{dd,max}(\Theta_c) = \frac{V_{dd}}{R_L} \cdot \frac{2\pi}{2\Theta_c - \sin(2\Theta_c)}$$ (3.20)

Based on equation for the drain current (3.13) and the equation for $I_{dd,max}$ equation (3.22) the maximum drain peak current can be calculated:

$$I_d\left(\frac{\pi}{2}\right) = I_{dd}\sin\left(\frac{\pi}{2}\right) - I_{dq} = I_{dd} - I_{dq}$$

$$I_{dd} = I_{dd,max}(\Theta_c) = \frac{V_{dd}}{R_L} \cdot \frac{2\pi}{2\Theta_c - \sin(2\Theta_c)}$$ (3.21)

$$\Rightarrow$$

$$I_{d,peak} = \frac{V_{dd}}{R_L} \cdot \frac{2\pi}{2\Theta_c - \sin(2\Theta_c)} - I_{dq}$$

Now $I_{ac}$ in equation (3.16) is expressed by $\alpha$ which is $I_{dd}$ relative to it maximum value $I_{dd,max}$:

$$I_{ac}(\alpha) = \alpha \cdot I_{dd,max}(\Theta_c) \cdot \frac{2\Theta_c - \sin(2\Theta_c)}{2\pi} = \alpha \cdot \frac{V_{dd}}{R_L}$$ (3.22)

So alpha is limited to value between zero and one.

$$\alpha = \frac{I_{dd}}{I_{dd,max}(\Theta_c)} \quad 0 \leq \alpha \leq 1$$ (3.23)

The drain DC current and the power consumption is calculated:

$$I_{DC}(I_{dd},\Theta_c) = \frac{1}{2\pi} \cdot \int_0^{2\pi} I_d(y)dy$$

$$= \frac{1}{2\pi} \cdot \int_{\pi/2 - \Theta}^{\pi/2 + \Theta} (I_{dd}\sin(y) - I_{dq}(\Theta_c))dy$$ (3.24)

$$= \frac{I_{dd}\sin(\Theta_c) - I_{dq}(\Theta_c) \cdot \Theta_c}{\pi}$$

$$= \frac{I_{dd}}{\pi} (\sin(\Theta_c) - \cos(\Theta_c) \cdot \Theta_c)$$
Again $I_{dd}$ is substituted with $\alpha$ in equation (3.23).

\[
I_{DC}(\alpha, \Theta_c) = \frac{I_{dd}}{\pi} \left( \sin(\Theta_c) - \cos(\Theta_c) \cdot \Theta_c \right)
\]

\[
= \frac{(\alpha I_{dd,\text{max}}(\Theta_c))}{\pi} \left( \sin(\Theta_c) - \cos(\Theta_c) \cdot \Theta_c \right)
\]

\[
= \frac{\alpha}{\pi} \left( \frac{V_{dd}}{R_L} \cdot \frac{2\pi}{2\Theta_c - \sin(2\Theta_c)} \right) \left( \sin(\Theta_c) - \cos(\Theta_c) \cdot \Theta_c \right)
\]

\[
= 2\alpha \frac{V_{dd}}{R_L} \cdot \frac{\sin(\Theta_c) - \cos(\Theta_c) \cdot \Theta_c}{2\Theta_c - \sin(2\Theta_c)}
\]

(3.25)

In Figure 3.8 the drain peak current is plotted against the conduction angle. It is seen that the peak current is increasing drastically for decreasing conduction angel. The current approaches infinity for the conduction angle approaching zero. So below a certain limit of the conduction angle it is not practical to realize a class C amplifier. The DC current increases a the conduction angle increases. This implies that the efficiency decreases because the AC current is constantly equal to $Vdd/R_L$.

![Figure 3.8: The drain peak current, eq. (3.21), current and DC current, eq. (3.25) normalized by ($V_{dd} = 1$ and $R_L = 1$). All values are for maximum output power ($\alpha = 1$) which implies that the AC current is equal to its maximum of one.](image)

Figure 3.9 is a zoom of Figure 3.8. There are added curves showing the parameters in of the drain current equation (3.13). Again we see that the same behaviour of drain peak current. But we also see the parameters of the drain current equation are increasing even more drastically for small conduction angles.

Another disadvantage of a small conduction angle is that the current pulse gets narrow which leads a high frequency spectral components. This can be seen in the two plots in Figure 3.10. To the left we see the harmonics of the drain currents plotted for a fixed peak drain current. This graph is the one typically shown in text books [5]. To the right the drain current is plotted for a fixed AC current (and output power) at the fundamental frequency similar to the other graphs shown in Figure 3.8 and Figure 3.9. Normally when a design is started the maximum power is given due to system requirements but the peak current is unspecified and can be adjusted for best performance.
The graphs Figure 3.10 shows that for small conduction angels (class C) a lot of energy lies in the high harmonics which means that the output transistor must have a very high bandwidth. It must accommodate the 5th harmonic without significantly attenuation. If we move to higher conduction angles (towards class A) the harmonics contents disappears completely which means that only the fundamental frequency components must be un-attenuated.

The power in the load depends only on the drive current $I_{dd}$ and the conduction angle.

$$P_{ac}(\alpha) = \frac{(I_{ac}(\alpha))^2 \cdot R_L}{2} = \alpha^2 \cdot \frac{V_{dd}^2}{2R_L}$$ (3.26)

$$P_{dc}(\alpha, \Theta_c) = I_{DC} \cdot V_{dd} = 2\alpha \frac{V_{dd}^2}{R_L} \frac{\sin(\Theta_c) - \cos(\Theta_c) \cdot \Theta_c}{2\Theta_c - \sin(2\Theta_c)}$$ (3.27)
The efficiency is calculated.

\[
\eta(\alpha, \Theta) = \frac{P_{ac, max}}{P_{DC, max}} = \frac{\alpha^2 \cdot \frac{V_{dd}^2}{2R_L}}{2\alpha \cdot \frac{V_{dd}^2}{R_L} \cdot \frac{\sin(\Theta_c) - \cos(\Theta_c) \cdot \Theta_c}{2\Theta_c - \sin(2\Theta_c)}}
\]

\[
= \alpha \cdot \frac{1}{4} \cdot \frac{2\Theta_c - \sin(2\Theta_c)}{\sin(\Theta_c) - \cos(\Theta_c) \cdot \Theta_c}
\]

(3.28)

Figure 3.11: Power and efficiency of the output stage as a function of the conduction angle.

A high power efficiency can be achieved for small conduction angles but unfortunately the design will become impractical due to high drain peak currents. In practice there will be a trade off between efficiency and a maximum allowable drain peak current.

### 3.1.4 Class F

The class F amplifier is similar to the class B amplifier except that the third harmonic is blocked using a parallel resonator \(C_2L_2\), see Figure 3.12. The third harmonic is according Figure 3.10 zero in class B operation. Therefore a third harmonic component must be generated changing the conduction angle slightly or by introducing clipping. If it is possible to set the amplitude of the third harmonic of the drain voltage to a 1/9 of the fundamental, the drain voltage will be flat at its maximum and approximate a square wave. This improves the efficiency from 78.5% to 88.4% [6].

### 3.1.5 Class D

Two transistors are used as switches such that either the positive or the negative voltage is connected to the output of the PA. The fundamental is lead through at passive resonant network to the load. There are ideally no losses in the PA since the switches are either on or off so the efficiency is 100% [6].
3.1.6 Class E
Is similar to class D but only 1 switch is used. The circuit complexity is lower than class C and a high efficiency is achieved[5]. It suffers from a high maximum drain voltage but is still very popular due to the high efficiency[7][8][9].

3.2 Characterisation of the PA
The power amplifier is characterised using the following measures.

- Load: The load impedance, resistive, under normal working conditions. The amplifier is realised for optimum performance (e.g. output power or efficiency) that are only guaranteed with the specific load.
- Pout: The output power delivered to the load.
- Power Gain: Specifies the ratio of the power at the input and the output of the amplifier. For small signals there are several definitions depending on the type of matching network and the load condition. The definitions are based on small signal two ports. The input of the port is connected a generator with finite output impedance, the source, and a load impedance called the load.
  - Operating gain: is defined as the ratio of the power delivered to the load and the power going into the port.
  - Available gain: is defined as the ratio of the available power at the output (of the PA) and the available power from the source.
  - Transducer gain: is defined as the power delivered to the load and the available power from the source.
- Supply voltage: Supply voltage for normal operation.
- Quiescent current: The current consumption. Is useful for calculation the total dissipated power.
- Efficiency: Two frequently measures are the drain efficiency and the power-added efficiency. The drain efficiency is the ratio between the power delivered to the load and the total dissipated power. The power added efficiency is the ration between the difference of the input and output power and the total dissipated power.
- $P_{1\text{-dB}}$: The one dB compression is the input power level where the output power is 1 dB lower than the level of an ideal amplifier with the same gain. The compression point is often determined by clipping effects.
- Operation frequency.
• Bandwidth: The bandwidth indicates the frequency range where the output power is constant within some specified deviations. The bandwidth is often limited due to a high Q matching network. The high Q is a result of a impedance transformation ratio between a small output impedance of the power transistor (1-10 Ohm) and the standard 50 Ohm used in all RF systems.

• The power control ensures a certain power level under different operation conditions. Chip-processing variations and temperature deviations influences the predictability of the output power, which necessitate regulation loop to adjust the power. The typically way to implement the control is to pick out at part of the output signal using a directional coupler. The signal is then rectified and low pass filtered and compared with a reference signal to determine the error. The error is amplified and fed back to a control input of the PA, normal called \( V_{PC} \). The \( V_{PC} \) controls the output power. The realisation of the power control within the PA can be carried out controlling the bias point of the output transistor. This conform to the fact that the power regulation of commercial amplifiers is highly non-linear. Similar to the AM/AM and AM/PM functions new functions describing the relationship between amplitude and phase of the PA output signal and the power control voltage \( V_{PC} \) can be defined.

• AM/AM conversion function is the relation between the input and the output amplitude.
• AM/PM: specifies the phase deviation at the output signal sweeping the input amplitude.
Chapter 4

Linearisation Techniques

Linearisation techniques are used to enhance the linearity of the Power Amplifier (PA) and thereby avoid inband distortion and adjacent band interference. Typically, the linearisation techniques are used in conjunction with amplification of amplitude modulated signal, such as QAM (Quadrature Amplitude Modulation), because the PAs distorts the envelope signal of the fundamental frequency. The higher harmonics of the output signal are normally not considered as they are removed by a low-pass filter. Many wireless systems have non or insignificant distortion caused by phase changes of the input signal. The reason for this is that the ratio of the bandwidth and the carrier frequency is much smaller than one, and the PA practically gives a constant group delay for all the channel frequencies. This is also reason that simplifications in predistortion methods (mapping simplifies to complex gain method) can be applied.

With linearisation methods it is possible to employ a strong non-linear amplifier with high efficiency and still achieve linear behaviour of the system. The additional power consumed by the linearisation system should be considered in the calculation of the overall efficiency of the system.

Table 4.1: Survey of linearisation techniques

<table>
<thead>
<tr>
<th>Linearisation</th>
<th>Analog</th>
<th>Digital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compensations techniques</td>
<td>Cartesian Loop</td>
<td>Mapping techniques</td>
</tr>
<tr>
<td></td>
<td>Polar Loop</td>
<td>Complex Gain Predistortion</td>
</tr>
<tr>
<td></td>
<td>Envelope feedback</td>
<td>Polar Predistortion</td>
</tr>
<tr>
<td></td>
<td>Phase Correcting FB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Feed Forward</td>
<td></td>
</tr>
<tr>
<td>Signal splitting techniques</td>
<td>CALLUM, LINC, Envelope Elimination and Restoration</td>
<td>LINC</td>
</tr>
</tbody>
</table>

Several linearisation techniques are available. In Table 4.1 the techniques are characterised by two different properties. The first property distinguishes whether the technique is analog or digital. The other, whether the technique is a compensation techniques or signal splitting technique (see “The LINC Techniques” on Page 46). The method which splits the signal into 2 constant envelope signals is insensitive to non-linearities in the amplifiers (LINC and CALLUM). The
rest of the techniques compensates for the non-linearity. Among them are feed forward, feedback and predistortion techniques. In the following the different techniques are presented.

4.1 The LINC Techniques

The LINC technique [10][11] was presented by Donald C. Cox in 1974. LINC is a abbreviation for “Linear Amplification with nonlinear components”. It splits the signal to be amplified, \( s \), into two constant envelope signals. The two signals are constructed in such way that the sum of the signals yields the original signal. The two signals are amplified using normal nonlinear amplifiers. Due to the constant envelope neither AM to PM conversion nor AM to AM conversion occurs. Summing the two amplified signals yields the original signal. The harmonic components originating from the nonlinear amplification are ignored as they are eliminated by a low-pass filter.

The LINC amplification is based on the following relation. The input signal is amplitude and phase modulated:

\[
\tilde{s}(t) = A(t) \cdot e^{i \varphi(t)}
\]

This signal is described by two signals with constant amplitude:

\[
s_1(t) = \frac{a_{\text{max}}}{2} \cdot \cos[2\pi ft + \alpha(t)]
\]

\[
s_2(t) = \frac{a_{\text{max}}}{2} \cdot \cos[2\pi ft - \alpha(t)]
\]

Where \( a_{\text{max}} \) is the maximum amplitude of \( A(t) \). \( \alpha(t) \) is the angle defined by

\[
\alpha(t) = \arccos \left( \frac{A(t)}{a_{\text{max}}} \right)
\]

The sum of \( s_1(t) \) and \( s_2(t) \) yields the original signal \( s(t) \)

![Figure 4.1: Linearisation using the LINC principle.](image)
An efficient method to calculate the constant envelope signals is best formulated using complex envelope signals (see Figure 4.1). First the vector \( \hat{e}(t) \) is calculated which is a scaled and rotated version of the input signal.

\[
\hat{e}(t) = j \cdot s(t) \cdot \frac{1}{A [s(t)]^2} - 1 \quad \tilde{s}(t) = A(t) \cdot e^{j\phi(t)} \tag{4.4}
\]

Then the constant envelope signals are calculated as the sum and the difference of the input signal, \( \tilde{s}(t) \) and the vector \( \hat{e}(t) \). This method is used in [12] to calculate the constant envelope signals \( \tilde{s}_1(t) \) and \( \tilde{s}_2(t) \). The calculation of \( \hat{e}(t) \) was realised as a table look up in a DSP to increase speed.

One disadvantage of the LINC system is the power combining [13]. If the combined signals are uncorrelated the insertion loss is 3dB which degrade the power efficiency significantly[14][15]. Although it is claimed that nearly 100% power efficiency can be achieved it has not yet been shown. To achieve high efficiency it is required that the PA’s is insensitive to variations in the load impedance[16][17][18].

Most work carried out on the LINC technique have been focusing on using a DSP [14][19][15][20] to ensure a high precision when calculating the constant envelope signal. But recently good results have been achieved using analog signal processing integrated on a chip [21]. This reassembles the first work [10] carried out on the LINC system but in a more advanced technology. The principle is in both cases based on feedback to calculate the two signal components.

### 4.2 CALLUM

The Combined Analogue Locked Loop Universal Modulator (CALLUM) [22] uses the same principal as the LINC amplifier in the sense that two constant envelope signals are combined to produce an amplitude and phase modulated signal. The two signals are created using a feedback loop instead of signal separation inside a DSP. The output of the PA is demodulated and each IQ branch are compared with the respective I and Q signal of the IQ-modulator to calculate the loop error. The high loop gain, provided by the amplifiers, ensures a small error. Thereby the demodulated IQ signal will approximate the IQ signals from the IQ modulator. The CALLUM system also have a problem combining the constant envelop signals with wasting 3dB power.

![Figure 4.2: CALLUM Linearisation.](image-url)
4.3 Cartesian Modulation Feedback

The Cartesian Modulation Feedback Method [11] is based on two feedback loops. One for the inphase (I) and one for the quadrature (Q) signal (see Figure 4.3). The output of the PA is demodulated using a quadrature demodulator. The demodulated I and Q signals are compared with the input I and Q signals and the error is calculated. The error is amplified, lowpass filtered, upconverted and led into the PA. The PA is controlled by the amplified error. If the loop gain is high then error will be small. This implies that the downconverted PA signal contains the same IQ components as the input signals. Thereby the linearisation is achieved.

The bandwidth is limited by the delays in the loop. The PA, the mixer and the filters all contribute with delays that limits the bandwidth of stable operation [23]. One way to compensate for the delay is to introduce a phase shift between the LO of the up- and down-conversion mixers.

Figure 4.3: Cartesian modulation feedback system.

The linearity of the up-converter is in principle not a limiting factor for the linearity because it can be perceived as a part of the non-linearity of the PA can therefore be compensated by the linearisation system. The phase imbalance of the IQ-up-converter is compensated too, but non-linearity and phase imbalance in the down-converter limits the achievable performance.

The design requirements of the down-converter in a linearisation system are different from the them in a receiver. In a linearisation system the input level is high and well defined. Therefore noise is expected to be a smaller problem whereas linearity is more severe. Mixer concepts such as the passive mixer-bridge with a buffer amplifier [24] invented in Michal Steyeart’s group are suitable because of the high linearity. The mixer is noisy if the “switches” are driven as voltage controlled resistors. But it is not expected to be a problem in a linearisation system due to the large PA signal.

4.4 Polar Modulation Feedback Loop

The polar modulation feedback system [25] (see Figure 4.4) uses the magnitude and the phase of the PA output signal as feedback signals opposed to the Cartesian modulation feedback which uses the inphase and quadrature signals. It seems more appropriate to use the envelope signal as one of the feedback signals because the distortion (AM/AM and AM/PM) is directly related to this signal. It is therefore more simple to track distortion and compensate it as we will see in Chapter 5.
The polar modulation feedback loop operates in the following way. First the output from the PA is downconverted to an intermediate frequency. Then the envelope of the signal is extracted. This is carried out by mixing the signal with an amplitude limited version of itself. The amplitude is also extracted of the input signal. The amplitude of the input signal and the PA signal is subtracted to calculate the amplitude error. This signal is led into the power regulation of the PA where the amplitude is adjusted. The phase error is calculated by the phase detector and used to adjust the phase of the VCO. In this way the phase distortion of the PA is compensated by a phase change in the VCO. Although the polar loop originally was based on a PLL loop with a VCO and a phase detector, the system in [28] (see Figure 4.5) must be characterised as a polar modulation feedback loop too because the feedback signals are in polar form. The phase is linearised by a phase shifter preceding the PA. The phase detector controls the phase shifter. Using a phase shifter makes it possible to use a simple phase detector opposed to a phase-frequency detector. The PA input signal in Figure 4.4 is a constant envelope signal whereas the corresponding signal in Figure 4.5 is a varying envelope signal. This means that the full amplitude modulation of the PA output signal is controlled by the error signal in Figure 4.4 whereas the error signal in
Figure 4.5: only compensates for amplitude deviations caused by the varying PA gain. The latter method is called multiplicative feedback [27], and the former one is additive feedback.

### 4.5 Envelope Feedback

Envelope feedback [26] is a special case of the polar modulation feedback where only the envelope of the signal is measured and corrected. In the simple case, the envelope detector consists of a diode as rectifier followed by a RC lowpass filter. The envelope feedback method can be used if amplitude distortion is the main cause of distortion. Envelope feedback is also used in conjunction with constant envelope modulation. For instance in GSM this method is used to control the output power, and turn on and off the power (ramping up or down). The envelope generator should be realized in the DSP together with the IQ-generator as considered in [26]. The envelope signals from the DSP should be D/A converted and compared with the envelope of the PA. In Figure 4.6 the power is regulated using a variable gain control (VGA) proceeding the PA. In other cases the lowpass filtered signal is led directly into the power control of the PA. The lowpass filter is necessary to ensure stability. Since the amplitude information is on both inputs of the VGA this is multiplicative feedback as explained in Section 4.4.

The feedback can be based on the power of the PA output signal instead of the envelope. In this case the principle is called power feedback [27].

**Figure 4.6: Linearisation using the envelope feedback method.**

### 4.6 Phase Correcting Feedback

Phase Correction Feedback [9] is another special case of the polar modulation feedback loop where only the phase of the signal is measured and corrected (see Figure 4.7). The output of the PA is led through a limiting amplifier to extract the phase information. Thereafter it is compared

**Figure 4.7: Phase correcting feedback system.**
with the phase information of the input signal. The phase difference calculated by the phase detector is used to control a phase shifter and thereby compensate the phase error at the output of the PA. In [9] the principle was used with a class E amplifier and the phase error was reduced from 30 degrees to 4 degrees. In [28] the phase difference was extracted using a branch line coupler followed by diode detectors and a difference amplifier. Similar to polar modulation feedback the phase detector do not need to be a phase-frequency detector when a phase shifter is employed for phase compensation.

### 4.7 Feed Forward Linearisation

The Feed Forward Linearisation techniques uses, similar to the negative feedback principle, an error signal to compensate the error. The error is calculated as the difference between the input signal and the attenuated output signal. This error is amplified and subtracted from the output of the PA. It may sometimes be necessary to insert delays to ensure that the signals are in phase before a subtraction. The advantage of the linearisation method is that a high efficiency non-linear amplifier can be used to provide most of the needed power, while a linear amplifier is used to amplify the error signal. The amplified error signal is relative small which means that relative little power is dissipated in it, although it is a linear amplifier. Feed forward is unconditional stable as no feedback is used. The critically aspect in this concepts is to match the gain and the delays. A calibration would not suffice as the gain and the delays are functions of the temperature. The method could however be used with some kind of adaptation [29][30] to ensure correct gain and delay time under all operation conditions.

![Diagram of Feed Forward Linearisation](image)

**Figure 4.8: Feed Forward linearisation.**

### 4.8 Digital Predistortion

In the following, three digital predistortion methods are presented: Mapping Predistortion, Polar Predistortion, Complex gain. The methods are distinguished by the number and the dimension of the Look Up Tables (LUT), which signal is used to address the LUT (IQ, magnitude, power) and adaptation speed. The clock frequency is typically 8-16 times higher than the channel bandwidth. Below a digital predistortion system is depicted. Apart from the typical blocks in the up-conversion path (LO, IQ-modulator and PA) some analog blocks are needed to deliver the feedback signal to the DSP. A quadrature demodulator converts the PA output signal to baseband. Two anti-aliasing filters and A/D converters convert the analog IQ signal to digital signals for the DSP. On-line correction of the predistortion table (the LUT) is based on the feedback of the PA. The modulated IQ-signal is transformed (polar, power) and quantized and used to address the LUT. The LUT value is used to modify the IQ signal before it is sent to the D/A converter. The actually IQ signal of the PA output is used to correct the table entries if needed. It is always
possible to achieve stable operation of a digital predistortion system because the sample rate in the feedback path and the adaptation speed of the LUT can be as slow as necessary. The sample rate can be chosen independent of the sample rate in the feedforward signal path.

### 4.8.1 Mapping Predistortion

Mapping Predistortion is a digital predistortion method based on a 2 dimensional Look Up Table (LUT). Each IQ pair from the modulator is predistorted using the LUT before it is led to the D/A converters. The output of the PA is downconverted, A/D converted and fed into the DSP. Then the IQ error is calculated and the corresponding entry in the LUT is updated based the adaptation algorithm. The advantage of mapping predistortion is that misalignments and nonlinearities in the conversion process, e.g. the quadrature up-conversion mixer, are cancelled [14]. The disadvantage is that the LUT is two dimensional \((I_{PD}, Q_{PD}) = f_{LUT}(I, Q)\) which results in many entries and a long adaptation time as a result. Further phase adjustment in the loop is needed to ensure an effective operation. This method can be perceived as the digital counterpart of the Cartesian Modulation Feedback. They are both able to correct the phase imbalance and the IQ up-conversion mixer (see Appendix A) and they both are sensitive to delays and phase difference between the LO used for the up- and down-conversion.
4.8.2 Polar Predistortion

Polar predistortion is based on two one-dimensional tables containing corrections of the magnitude and the phase. The Look Up Table (LUT) is addressed using the length of the IQ vector coming from the IQ modulator. First the magnitude is corrected, then the phase is corrected using the new magnitude to address the phase table. Two additional rectangular to polar conversions are needed compared mapping predistortion, but the adaptation time is significantly shorter than in the case of mapping predistortion due to a smaller number of entries in the two LUTs.

![Digital Polar Predistortion](image)

**Figure 4.11: Digital Polar Predistortion.**

4.8.3 Complex Gain Predistortion

Complex Gain Predistortion[14][31] is similar to Polar predistortion, except that the two LUTs are addressed using the quantized power of the IQ signal and the correction is carried out multiplying a complex gain factor to the IQ-modulated signal. No rectangular to polar conversion is needed as in polar predistortion, but the power, $I^2 + Q^2$, has to be calculated which is less computational load compared to rectangular to polar conversion, $\sqrt{I^2 + Q^2}$ and $\text{atan}(Q/I)$.

![Digital Complex Gain Predistortion](image)

**Figure 4.12: Digital Complex Gain Predistortion.**
4.8.4 Polynomial predistortion

In Polynomial predistortion[32][34] the predistorted signal is calculated using a complex polynomial. The IQ signal is inserted in the polynomial to calculate the predistorted signal. This requires many complex multiplications, but no memory is needed as no look up table is used. Also the adaptation algorithm requires high computational power.

4.9 Remarks on PA isolation and sensitivity to external disturbance

Disturbances received on the antenna from other mobile phones in the transmit band could disturb the envelope feedback or power feedback linearisation system.

One solution to this problem is to insert an isolator after the PA (see Figure 4.13). This will attenuate the disturbers. Further it will reduce the influence of a varying antenna impedance. An isolator has an insertion loss of typically 0.3-0.5dB.

Figure 4.13: Isolator avoids disturbance from other transmitters received by the antenna.

Another method to reduce the effect of disturbers (e.g. tv broadcasting) is to insert a filter in the feedback path (see Figure 4.14). It is an advantage that no filter is in the signal path between the PA and the antenna which would lead to reduced power efficiency for the transmitter stage.

Figure 4.14: A BP-filter in the FB reduces out-of-band disturbances received by the antenna.
Chapter 5

Analysis and design of Linearisation Systems

In this chapter the Cartesian Modulation Feedback system and the Polar Modulation Feedback system will be analysed. The main focus will be on Polar loop but the Cartesian loop will be analysed first and later compared to the polar loop linearisation system. The basic non-linear equation for settlement of the feedback loop is found and the difference between Cartesian loop and Polar loop is discussed. The overall system distortions function for the linearisation system is found (AM/AM and AM/PM) based on the distortions functions of the power amplifier. The dynamics of the polar loop is analysed and design guidelines for the amplitude loop are given. An interesting recognition is that in a polar linearisation system the envelope and phase loop can be solved separately whereas in the Cartesian loop the feedback loops are tight coupled due the fact that the distortion of the power amplifier is related to the envelope of the input signal only. This gives some advantages when the output signal of the linearisation system is simulated. These advantages are exploited in the simulation guidelines given in this chapter. The fact that the PA distortion is related to the envelope makes polar loop the better choice for handling non-linear power amplifiers. The amplitude can be used as a state variable to monitor the loop gain. It is shown that the loop bandwidth is a function of the varying PA gain and that this leads to a trade off between precision and loop stability. To solve that problem a transmitter architecture is suggested alleviated the varying loop gain problem without violating the stability of the loop. Simulations are shown that verifies the improvements of the architecture.

At the end of the chapter examples will be given on how non-linearities affect the signal quality of a modulated signal. From the examples it will be shown that the spectral leakage (ACP) is the limiting requirement compared to inband signal quality (error vector magnitude) given typical radio specification. In some publications there have been argued for that envelope linearisation is more important than phase linearisation [26] which could be the case for certain PA’s and applications. From the examples it is shown that amplitude as well as phase distortion degrades the signal quality. So if a general and robust linearisation system should be designed both envelope and phase feedback should be employed.

The work presented in this chapter has lead to a publication[35] and that two patents has been filed [36],[37].
5.1 Amplitude linearity of Cartesian Modulation Feed Back System

The distortion in a Cartesian modulation feedback linearisation system is analysed. The analysis is based on the block diagram Figure 5.1. It shows the signal processing carried out. All signals are complex signals and complex envelope signals. The signals $s_1$ and $s_2$ are complex time continuous baseband signals representing the inphase signal, the quadrature signal. $e_1$ is the error between the input signal to the system and the down converted PA signal. The signals $s_3$ and $s_4$ are pre-envelope of the RF signals. The signals $\tilde{s}_3$ and $\tilde{s}_4$ are the respective complex envelope signals. Based on the block diagram we can write the following equations.

The input signal is described by the inphase and quadrature signal.

$$s_1 = I(t) + j \cdot Q(t)$$  \hspace{1cm} (5.1)

The error is the difference between the input signal and the down-converted PA signal.

$$e_1 = s_1 - s_2$$  \hspace{1cm} (5.2)

The input signal to the PA is the amplified an up-converted error.

$$s_3 = e^{j\omega t} \cdot k \cdot e_1 \quad \tilde{s}_3 = k \cdot e_1$$  \hspace{1cm} (5.3)

The gain stage with value $k$ also represents a filter in the sense that the value can be complex corresponding to a filter value at steady state.

The output of the PA is in the following modelled by the small signal gain, $G$, the gain ideality factor, $A$ and the phase distortion, $\Theta$.

$$s_4 = GA(\|s_3\|) e^{j\Theta(\|s_3\|)} \cdot s_3 \quad \tilde{s}_4 = GA(\|\tilde{s}_3\|) e^{j\Theta(\|\tilde{s}_3\|)} \cdot \tilde{s}_3$$  \hspace{1cm} (5.4)

Notice that gain ideality factor and the phase distortion are functions of the amplitude of the input signal, $s_3$, of the PA because the amplitude-distortion and the phase distortion of a narrow band PA are functions only of the input amplitude of the PA. There is a simple relation between these functions and the AM/AM and AM/PM as it will be shown later.

The PA signal is attenuated and down-converted and used as feedback signal.

$$s_2 = \alpha \cdot s_4 e^{-j\omega t}, \quad 0 < \alpha \leq 1$$  \hspace{1cm} (5.5)
Combining the equations (5.1)-(5.5) the downconverted PA signal, $s_2$ is obtained as a function of the input signal, gain ideality factor and phase distortion:

$$s_2 = e^{-j\omega t} \cdot \alpha GA(|s_3|) e^{i\Theta(|s_4|)} \cdot e^{i\omega t} \cdot (s_1 - s_2)k$$

$$\Rightarrow s_2 = \frac{\alpha GA(|s_3|) e^{i\Theta(|s_4|)} \cdot k}{1 + \alpha GA(|s_3|) e^{i\Theta}} \cdot s_1$$  \hspace{1cm} (5.6)

Now the other signals in the system can be expressed by the input signal, gain ideality factor and phase distortion. The error is:

$$e_1 = s_1 - s_2 = \frac{1}{1 + \alpha GA(|s_3|) e^{i\Theta(|s_4|)} \cdot k} \cdot s_1$$  \hspace{1cm} (5.7)

The PA input signal.

$$s_3 = e^{i\omega t} \cdot k \cdot e_1 = \frac{k}{1 + \alpha GA(|s_3|) e^{i\Theta(|s_4|)} \cdot k} \cdot e^{i\omega t} \cdot s_1$$  \hspace{1cm} (5.8)

The complex envelope is the pre-envelope translated down in frequency by a reference frequency corresponding to the carrier frequency. We also use that the magnitude of the pre-envelope is equal to the magnitude of the complex envelope.

$$\tilde{s}_3 = s_3 e^{-j\omega t} = \frac{k}{1 + \alpha GA(|s_3|) e^{i\Theta(|s_4|)} \cdot k} \cdot s_1$$  \hspace{1cm} (5.9)

The PA output signal is:

$$\tilde{s}_4 = GA(|\tilde{s}_3|) e^{i\Theta(|\tilde{s}_4|)} \cdot \tilde{s}_3 = \frac{k\alpha GA(|\tilde{s}_3|) e^{i\Theta(|\tilde{s}_4|)}}{1 + \alpha GA(|\tilde{s}_3|) e^{i\Theta(|\tilde{s}_4|)} \cdot k} \cdot s_1$$  \hspace{1cm} (5.10)

Now the amplitude of $s_3$ is calculated using equation (5.9).

$$|\tilde{s}_3| = \left| \frac{k}{1 + \alpha GA(|\tilde{s}_3|) e^{i\Theta(|\tilde{s}_4|)} \cdot k} \right| \cdot |s_1|$$  \hspace{1cm} (5.11)

The equation can be written based on real numbers.

$$|\tilde{s}_3| = \frac{|k| \cdot |s_1|}{\sqrt{1 + 2\alpha GA(|\tilde{s}_3|) \cos [\Theta(|\tilde{s}_3|) + \arg(k)]|k| + [\alpha GA(|\tilde{s}_3|)|k|]^2}}$$  \hspace{1cm} (5.12)

Assume that gain ideality factor $A(|\tilde{s}_3|)$ and phase distortion $\Theta(|\tilde{s}_3|)$ are known functions. Then equation (5.11) is one equation with one unknown variable namely the magnitude of $s_3$. So if that equation can be solved we have the magnitude of the PA input signal $s_3$. This is all we need to calculated all other signals in the system in Figure 5.1 because equation (5.6) to equation (5.10) only depend on the input signal $s_1$ and the magnitude of $s_3$. The gain ideality factor $A(|\tilde{s}_3|)$ and phase distortion $\Theta(|\tilde{s}_3|)$ can be related directly to the AM/AM and AM/PM characteristic as it will be shown below. But first let us make some comments on derived equations:
To calculate the magnitude of input signal of the PA ($s_3$) one non-linear equation with one unknown has to be solved.

All signals in a Cartesian Modulation Feedback loop will be known if the magnitude of the the PA input signal ($s_3$) and the input-signal to the system ($s_1$) is known.

The AM/AM and the AM/PM characteristic of the system (Cartesian Modulation Feedback system) can be found by equation (5.10) if we can find the magnitude of the PA input signal ($s_3$) for every input-signal to the system ($s_1$).

So if we can solve the non-linear equation (5.11) we are able to abstract one level in the sense that we can calculate the AM/AM and AM/PM characteristic of the system and use them to characterize the whole linearisation system as a black box. The above statements assume that the gain stage amplifies any frequency by the gain $k$. This is of course not the case in a real feedback system because a band limitation is needed to guarantee stability. But if the bandwidth is large enough (normally 2-5 times larger the bandwidth of the modulated input signal) the above statements will hold.

As mentioned the non-linearity of a power amplifier is specified through AM/AM characteristic and the AM/PM characteristic. To relate these functions directly to gain ideality factor and phase distortion function used in the previous we calculate the relations by using the definitions. The magnitude of the PA output can be expressed by equation (5.4) and by the AM/AM characteristic.

$$|s_4| = GA(|s_3|) \cdot |s_3| = AMAM(|s_3|)$$

$$GA(|s_3|) \cdot |s_3| = AMAM(|s_3|)$$

The phase relation can be calculated the following way.

$$\arg(s_4) = \Theta(|s_3|) + \arg(s_3) = AMPM(|s_3|) + \arg(s_3)$$

$$\Theta(|s_3|) = AMPM(|s_3|)$$

The argument of $s_3$ cancels out so the phase distortion is equal to the AM/PM characteristic. The AM/AM and AM/PM characteristic of the entire system can as mentioned be found by equation (5.10). The functions are however slightly different than we are used to because the input signal is a complex number representing a low frequency signal. That is different from the functions representing the PA where both signals represents RF signals. The AM/AM characteristic is the amplitude of the output of the system $|s_4|$ as a function of the magnitude of the input to the system $|s_1|$.

$$AMAM_{sys}(|s_1|) = |s_4| = \left| \frac{k \alpha GA(|s_3|) e^{j\Theta(|s_3|)}}{1 + \alpha GA(|s_3|) e^{j\Theta(|s_3|)} \cdot k} \cdot |s_1| \right|$$

Remember that amplitude of $s_3$ is function of magnitude of $s_1$ (see equation (5.11)) so the AM/AM characteristic is fully defined by magnitude of $s_1$. 
The AM/PM characteristic is the phase difference between the input signal and the output signal of the system.

\[
\arg(\tilde{s}_4) = \arg\left( \frac{k\alpha GA(\tilde{s}_3)}{1 + \alpha GA(\tilde{s}_3)} e^{j\Theta(\tilde{s}_3)} \cdot k \right) + \arg(s_1) \Rightarrow
\]

\[
AMPM_{sys}(s_1) = \arg(\tilde{s}_4) - \arg(s_1) = \arg\left( \frac{k\alpha GA(\tilde{s}_3)}{1 + \alpha GA(\tilde{s}_3)} e^{j\Theta(\tilde{s}_3)} \cdot k \right)
\]

(5.16)

Again the amplitude of \( s_3 \) is function of magnitude of \( s_I \) (see equation (5.11)) so the AM/AM characteristic is fully defined by the input signal \( s_I \).

Below an example of the derived equations is given. Assume that we want to calculate the linearity of the linearisation system with the AM/AM and AM/PM characteristic of the PA shown in Figure 5.2. The used loop gain is 100 (k=100).

![Figure 5.2: AM/AM and AM/PM characteristic of a PA.](image)

The overall characteristic of the system can be calculated by sweeping the input amplitude and solving equation (5.11) and (5.10). The result of such a sweep is shown in Figure 5.3. Also gain ideality factor is shown. It is seen that for increasing input amplitudes the PA compresses. At the same time the input amplitude to the PA is increased by the system to compensate for the
PA signal compression. Thereby a more linear system is achieved. The relative error is shown in Figure 5.4. It is seen that the relative amplitude error increases with higher amplitudes. This is because the PA lowers the loop gain which result in a higher relative error.

In Table 5.1 a comparison for to different loop gains are shown. It is seen that the linearity is improved with the loop gain. Further it is noticed that the error in a loop is proportional to one divided by the loop gain. This relation also recognize here. The reason that it is not exact is due to the fact that the gain of the PA modifies the loop gain. The linearity definition used for the comparison is the normalised RMS error between the curve and a line fitted to the curve.

<table>
<thead>
<tr>
<th>Loop Gain</th>
<th>Linearity of Linearised system</th>
<th>Linearity of PA alone</th>
<th>Improvement (ratio)</th>
</tr>
</thead>
<tbody>
<tr>
<td>k=10</td>
<td>21.3e-6</td>
<td>274e-6</td>
<td>13</td>
</tr>
<tr>
<td>k=100</td>
<td>3.37e-6</td>
<td>274e-6</td>
<td>81</td>
</tr>
</tbody>
</table>

5.2 Polar modulation feedback

In the following a new transmitter architecture for polar modulation feedback linearisation of RF power amplifiers is presented. The architecture is adapted to digital transmitters which makes it possible to move some of the analog signal processing to the digital domain. An improvement of the varying loop-gain problem in feedback linearisation systems, that is easily incorporated in the new architecture, is presented. This improvement increases the stability margin and reduces the variations of the bandwidth which leads to higher linearity of the system. The presented architectures are suited for monolithic integration [38]. Simulations documenting the improvements and showing important design constraints are presented. System simulations are based on the GSM/EDGE system [39] and a measured power amplifier (PA).

5.2.1 Polar linearisation in a digital transmitter.

A polar modulation feedback linearisation system ([38],[25]) is shown in Figure 5.5. The system is not optimal for modern digital transceivers because the essential signals (amplitude and phase) are derived by analog circuits at an intermediate frequency.
We suggest two alternative architectures that generates the essential signals (amplitude and phase) digitally in the DSP and allow us to eliminate a limiting amplifier and a mixer used for the amplitude extraction. Now the interface between the DSP and the RF-frontend is in polar coordinates opposed to the traditional cartesian (inphase and quadrature) interface.

5.2.1.1 Polar loop with an RF-phase shifter

The first polar loop architecture is reminiscent of the envelope elimination and restoration techniques in the sense that the phase and the amplitude is dealt with separately (Figure 5.6). Instead of eliminate the envelope, the DSP delivers the desired envelope and a constant envelope signal. No amplitude elimination is needed. The phase signal modulates the LO signal. The signal is used as a reference to the phase feedback loop. The phase feedback loop consist of two limiters, a phase detector and a phase shifter. The phase detector measures the phase error between the output of the PA and the reference signal. This signal is used to control the phase shifter. The

Figure 5.5: Traditional polar modulation feedback.

Figure 5.6: Alternative amplitude and phase linearisation.
phase shifter thereby compensates for the phase error generated by the PA. The carrier frequency is determined by the LO. The architecture can also be used with a frequency offset by inserting a mixer in the phase feedback loop. In this architecture the frequency synthesis is independent of the linearisation.

The amplitude of the wanted signal is dealt with in a separate loop. In this concept the output level of the PA is regulated directly on the power control of the PA. It could, however, also be realised by using variable gain amplifier (VGA) in front of the PA, or by modulating the supply voltage directly. Today switch mode power supply achieves 90% to 95% at DC but if the supply should track the varying amplitude the efficiency will degrade. The error signal is calculated as the difference between the reference amplitude from the DSP and the attenuated amplitude of the PA output signal. This concept is fully analog in the front-end, which makes it interesting compared to digital solution where power consuming A/D converters are used. It is possible that the whole front-end, the linearisation system and the PA, could be integrated on one chip. It should not be necessary to implement the front-end in a CMOS technology as no digital circuits, such as the DSP, are used.

5.2.1.2 Polar loop linearisation using a VCO

The second alternative concept is also an analog solution with base band interface in polar coordinates. The phase distortion of the PA is compensated by embedding the PA in a phase locked loop. The phase linearisation is combined with the frequency synthesis, which offers some advantages. It is possible to save functional blocks such as VCOs, if the same block can be used to the frequency synthesis and the linearisation. Further, this will save power. The power is controlled in the same manner as explained in the previous concept. The output frequency is given by the reference frequency, $f_{\text{ref}}$, times the division number of the frequency divider. The phase of the PA output signal is equal to the phase of the DSP times the division number. This phase scaling must be compensated in the DSP by a division so the phase of the modulated signal effectively is scaled by one.

**Figure 5.7: Combined frequency synthesis and linearisation.**
5.2.2 Analysis of the Polar loop

The system Figure 5.7 is now analysed further but first we modify the architecture to accommodate infinite phase deviations by representing the phase by an inphase and quadrature signal (see Figure 5.8). We also employ a mixer to convert the inphase and quadrature signal to an intermediate frequency, \( f_{\text{ref}} \). In this way we have also achieved an offset loop architecture (see Section 1.1.3 "Offset loop transmitter" on Page 6). A limiter and a mixer is eliminated of the original Polar Modulation architecture shown in Figure 5.5 which improves the linearity and reduces the circuit noise. Furthermore, Cartesian to polar conversion is handled more power efficient and precise in the DSP.

The functionality is explained more in detail now. The DSP generates an envelope signal, \( a(t) \) and two quadrature signals, \( I'(t) \) and \( Q'(t) \), representing a RF constant envelope signal. The quadrature signals are up-converted and used as reference signal for the phase detector. The necessary extra DSP calculations are as follows:

\[
A(t) = \sqrt{I(t)^2 + Q(t)^2}
\]

\[
I'(t) = \frac{I(t)}{A(t)} \quad Q'(t) = \frac{Q(t)}{A(t)}
\]

where \( I(t) \) and \( Q(t) \) are the digital modulated signals normally provided by the DSP. The calculations can easily be carried by a modern DSP, if necessary by using table look up techniques. The resulting output of the system Figure 5.8 is as follows:

\[
s(t) = A(t)(\text{Re}[(I'(t) + jQ'(t))\cdot e^{j\omega_c t}])
\]

which is the normal linear modulated signal. In the following the new system will be analysed. The carrier frequency is given by

\[
f_{\text{out}} = \frac{N}{M} f_{\text{ref}}
\]

Instead of a divider a down-conversion mixer can be used. In this way the phase scaling as discussed in the last section is avoided.
5.2.2.1 Loop settlement and linearity compared to Cartesian loop.

To make the comparison easy the Cartesian loop Figure 5.1 is modified to represent a envelope feedback system as shown Figure 5.9. The amplitude feedback loop is not affected by the phase loop and can be treated independently as we will see in the next chapter.

![Figure 5.9: Cartesian MFB modified to an envelope feedback system.](image)

The mixer used for down conversion in the feedback path is replaced by an envelope detector. This means the feedback signal \( s_2 \) is a real and positive number. Also the input to the feedback system is now a real number representing the amplitude. This implies that the error signal \( e_1 \) is a real number. The gain stage and the up-conversion mixer can be perceived as a variable gain amplifier (VGA). The loop gain factor \( k \) represents the gain of the loop filter. Equations for the system are as follows.

- The error is a real number.
  \[
  e_1 = s_1 - s_2 \tag{5.21}
  \]

- The input signal to the PA is the scaled LO signal.
  \[
  s_3 = e^{j\omega t} \cdot k \cdot e_1 \quad \tilde{s}_3 = k \cdot e_1 \tag{5.22}
  \]

- The output of the PA is unchanged.
  \[
  s_4 = GA(s_3) e^{j\Theta(|s_3|)} \cdot s_3 \quad \tilde{s}_4 = GA(|\tilde{s}_3|) e^{j\Theta(|\tilde{s}_3|)} \cdot \tilde{s}_3 \tag{5.23}
  \]

- But the feedback signal is the amplitude of the PA output. No phase information is present due to the envelope detector.
  \[
  s_2 = \alpha \cdot |\tilde{s}_4| = \alpha GA(|\tilde{s}_3|) \cdot |\tilde{s}_3| \quad 0 < \alpha \leq 1 \tag{5.24}
  \]

Solving the above equations for the PA input signal, \( s_3 \), yields the equation.

- \[
  \frac{\tilde{s}_3}{k} = s_1 - \alpha GA(|\tilde{s}_3|) \cdot |\tilde{s}_3| \tag{5.25}
  \]

The right side of the above equation is real numbered which implies that the left side must be too. So the phase of the PA input signal is determined by the phase of the loop filter \( k \).

- \[
  \arg(\tilde{s}_3) = \arg(k) + \arg(p) \tag{5.26}
  \]
where \( p \) is the sign of the right side of equation (5.25) which corresponds to the sign of the error \( e_1 \) in the system.

\[
p = \begin{cases} 
1 & \text{for } [s_1 - \alpha GA(\tilde{s}_3)] \cdot |\tilde{s}_3| \geq 0 \\
-1 & \text{otherwise}
\end{cases} 
\]  

(5.27)

The magnitude of the PA input signal \( s_3 \) can be calculated by calculating the magnitude on both sides of the equality sign of equation (5.25). We also rearrange the equation to make it look like the basic equation for Cartesian loop (equation (5.11) on Page 57).

\[
|\tilde{s}_3| = \frac{|k| \cdot p}{1 + \alpha GA(|\tilde{s}_3|)|k|} \cdot |s_1| = \frac{|k|}{1 + \alpha GA(|\tilde{s}_3|)|k|} \cdot |s_1| 
\]

(5.28)

The above equation is compared to the corresponding settling equation for the Cartesian loop in Table 5.2. The difference lies in the denominator \( 1 + L(|\tilde{s}_3|) \) of equation (5.28) and (5.11). The Cartesian loop can pursue all phases in order to settle. This affects the phase of error signal in such way that the PA output signal is compensated in phase as well as amplitude. It is seen that the error is small if the loop gain is kept high. This can be assured by making the gain of the gain stage \( k \) high (see equation (5.7)). The loop gain of the polar loop is limited to a real number due to the envelope detector which removes any carrier phase information. This corresponds to a solution in one dimension namely the amplitude. In case of Cartesian modulation feedback the error can be guaranteed low by keeping the loop gain high through \( k \). But the error signal in polar modulation feedback represents the amplitude opposed to Cartesian modulation where the error is in cartesian coordinates and therefore contains amplitude as well as phase. This also confirms that a phase feedback loop is necessary to make a complete linearisation in a polar modulation feedback linearisation system. This may seem as a penalty in hardware complexity of the polar loop but it should be noted that the hardware (subtraction circuit, gain stage and mixer) processes two-dimensional signals in Cartesian modulation feedback (the inphase and quadrature signal) whereas the envelope loop processes a one dimensional signal (the amplitude).

| System       | Settling equation | Loop Gain \( L(|\tilde{s}_3|) \) | Magnitude/ Argument \( |L(|\tilde{s}_3|)| \) |
|--------------|-------------------|---------------------------------|---------------------------------|
| Cartesian MFB | \[\tilde{s}_3 = \frac{k}{1 + \alpha GA(|\tilde{s}_3|) e^{j\Theta} |\tilde{s}_3|} \cdot |s_1|\] | \( \alpha GA(|\tilde{s}_3|) e^{j\Theta} |\tilde{s}_3| \cdot k \) | \( \alpha GA(|\tilde{s}_3|) |k| \) |
| Polar MFB    | \[\tilde{s}_3 = \frac{|k|}{1 + \alpha GA(|\tilde{s}_3|)|k|} \cdot |s_1|\] | \( \alpha GA(|\tilde{s}_3|) |k| \) | \( \alpha GA(|\tilde{s}_3|) |k| \) |

5.2.3 Dynamics of the polar loop linearisation system.

The analysis of the transmitter architecture in Figure 5.8 is based on the signal diagram shown in Figure 5.10. The PA is modelled by the small signal gain, the gain ideality factor and the
SECTION 5.2 - POLAR MODULATION FEEDBACK

phase distortion function. In the following the small signal gain and the gain ideality factor is represented by one factor, $G_{PA}$, for simplicity. The loop dynamics in Figure 5.10 is set by the loop filters $F(s)$ and $G_F(s)$. For a fixed PA-gain the amplitude and phase feedback loops are described with the following equations.

$$\frac{A_{PA}(s)}{A(s)} = \frac{G_{PA}G_F(s)}{1 + B G_{PA}G_F(s)} = H_{AL}(s)$$  \hspace{1cm} (5.29)

$$\Phi_{PA}(s) = H_{Ref}(s) \cdot \Theta_{Ref}(s) + H_{PM}(s) \cdot \Theta_{AMPM}(s)$$  \hspace{1cm} (5.30)

where $G_{PA}$ is the PA-gain
$B$ is the gain of the amplitude detector
$\Phi_{Ref}(s)$ is reference phase provided by the DSP
$H_{PM}(s)$ is the injected phase distortion due to amplitude to phase conversion (AM/PM).

Since the above equations are based on a linear system we had to assume that the PA gain was constant. This only holds if the input amplitude is constant, but if we assume that the amplitude variations are small, the equations will be good approximated. Furthermore the understanding of the equations can be difficult if the amplitude input signal contains two or more frequencies components. The problem is that superposition is in principle not possible because of the varying PA gain. In this case we have to assume that one of the frequency components are dominating and determines the time varying PA gain. Then the other contributions from spectral components can be calculated based on their respective frequency and the transfer functions (which are time varying). Even though the equations in principle only hold for a linear system, they make it possible to give statements on the dynamics and stability as we will see in the following. An overall linear system is achieved if $H_{AL}(s)$ and $H_{Ref}(s)$ are close to one and $H_{PM}(s)$ is close to zero in the band of interest ($f < f_{lin}$). This bandwidth should be chosen to 2-4 times the channel bandwidth, $f_{ch}$, to be able to attenuated spectral regrowth in the neighbour channels [14]. If we assume that a charge pump is employed within the phase loop, $H_{Ref}(s)$ as-

---

**Figure 5.10: Polar loop signal model**

---

66
sumes an ordinary type 2 PLL transfer function [40] and \( H_{PM}(s) \) assumes a VCO phase transfer function.

### 5.2.3.1 Simulating the polar loop

The polar loop feedback system Figure 5.10 can be simulated in a simple way due to the fact that there is only a coupling from the envelope loop to the phase loop and not in the other direction. This can be seen from the transfer functions (5.29) and (5.30). To calculate the system output the following steps should be carried out.

1) Generate the modulated signal as described in Section 1.3.1 on Page 11.
2) Convert the inphase and quadrature signal to an envelope signal and a phase signal.
3) Simulate the amplitude loop
4) Calculate the disturbance to the phase loop by applying the AM/PM function to the calculated amplitude of the PA-input signal.
5) Simulate the phase loop with two phase inputs signals. In the phase loop the phase of the modulated signal as well as the phase disturbance from the amplitude loop have to be accounted for in the simulation. The dynamics of the phase loop is unaffected by the varying PA gain. So, the loop can usually be linearised which implies that super-position can be used.
6) Combine the result of the simulation of the AM-loop with the result of phase-loop simulation to a complex envelope signal of the output.
7) Evaluate the signal quality as suggested Section 1.3.5 on Page 18.

It is an advantage that the phase and amplitude loop can be simulated separately because all the non-linear effects takes place in the amplitude loop due to the varying PA gain. So in many cases, a transient simulation or a harmonic balance simulation must be carried out for the amplitude loop. The phase loop can be considered as a linear system and therefore it would in many cases be sufficient to uses the Fast Fourier Transformation to calculate the output signal of the phase loop. The linear phase loop also allow super-position to be used for the two input signals (see equation (5.30)).

In Cartesian feedback system this kind of fragmented simulation is not possible because the phase distortion affects both the inphase and quadrature signal at the same time. Also the input to the AM/AM function and the AM/PM function is an amplitude signal which is a function of both quadrature signals.

### 5.2.3.2 Loop gain problem

The main problem of using feedback in linearisation systems is that the loop-gain decreases for a decreasing PA-gain. For our PA (Figure 5.11) the PA-gain approaches zero for small and large amplitudes due to compression and turn-off phenomenons. Therefore, the system can only be designed to operate properly in an amplitude interval, where a minimum PA-gain is ensured. For a higher order loop filter, \( G(s) \), the maximum filter gain is determined by a minimum stability gain margin, and by the maximum PA-gain, \( G_{PA,max} \). Therefore the maximum amplitude error \( (H_{AL}(s) - 1) \) cannot be minimized by choosing an arbitrarily high filter gain. Another performance limitation of the linearisation system is the delay between the phase, \( \Phi_{PA}(t) \), and the amplitude signal \( A_{PA}(t) \). The delay of the signals is determined by the closed loop transfer functions \( H_{AL}(s) \) and \( H_{Res}(s) \) in equation (5.29) and equation (5.30). A constant delay difference between these signals could be compensated in the DSP, but a varying delay would require significant DSP processing power. From equation 5.29 we see that the amplitude transfer func-
tion varies due to varying PA-gain. The phase transfer function does not depend on the PA-gain and represents a fully linear system.

The bandwidth of amplitude transfer function varies between the bandwidth of the loop filter, \( f_g \), and a bandwidth approximately \( L \) times larger. Where \( L \) is the loop-gain. To ensure a low variation of the group delay a minimum bandwidth of the closed loop system, \( H_{AL}(s) \), must be guaranteed. E.g. if a first order loop filter, \( G(s) \), with one pole at \( f_g \) and a DC gain of \( G_0 \) is employed, the closed loop system is of first order, with a minimum bandwidth depending on the filter pole and the minimum PA-gain.

\[
f_{CL, min} = (1 + BG_{PA, min}G_0)f_g
\]

The group delay is as follows

\[
\tau(f, f_{CL}) = f_{CL}/(2\pi f^2 + f_{CL}^2)
\]

The group delay is nearly independent of the input frequency of the amplitude signal \( A(t) \) if the closed loop bandwidth, \( f_{CL} \), is significantly larger than the maximum input frequency (\( f_{mod} \)).

\[
\tau(f) \equiv 1/(2\pi f_{CL}^2) \text{ for } f^2 \ll f_{CL}^2
\]

In this case the minimum and maximum group delay can be written as follows.

\[
\tau_{max} \equiv 1/(2\pi f_{CL, min}^2) \quad \tau_{min} \equiv 1/(2\pi f_{CL, max}^2)
\]

A constant group delay results in a delayed but undistorted amplitude signal. The RF signal will, however, be distorted because of the time offset between the amplitude and phase signal. The amount of distortion also depend on the type of modulation. To relate the delay and the distortion, we calculated the distortion as a function of the delay for a long GSM/EDGE modulated [39] signal (Figure 5.12). We found that the ACP (spectral leakage) requirements were more restrictive than the EVM (inband signal quality) requirements (EVM<9%). This is also in agreement the simulations in Section 6.7.4 and with the simulations on a QPSK signal in Section 5.3. In Figure 5.12 a graph is showing ACP at an offset of 200kHz and 400kHz. They are required to be below -30dB and -60dB [39] respectively. So the maximum allowable delay is 23.5ns corresponding to an ACP of -60dB at an offset of 400kHz. Based on a minimum PA-gain, equation (5.31) and equation (5.34) and the relation in Figure 5.12, an estimate for the minimum allow-
able filter bandwidth can be found. This is used to give a set of design guidelines for an envelope feedback loop.

5.2.3.3 Design guidelines for an envelope feedback system.

1) Choose the modulation bandwidth, $f_{\text{mod}}$. This bandwidth is normally determined by the radio standard.

2) Choose the minimum closed loop bandwidth, $f_{\text{CL,min}}$, of the feedback system in such a way that the following two requirements are fulfilled.
   2a) The minimum closed loop bandwidth is significantly higher than the modulation bandwidth, $f_{\text{mod}}$. This is to ensure a constant delay for all frequencies of interest.
   
   \[ \frac{f_{\text{mod}}^2}{f_{\text{CL,min}}^2} \ll 1 \]

   2b) The closed loop bandwidth, $f_{\text{CL,min}}$, must be large enough to ensure a ACP requirement specified by the radio standard. Use the graph similar to Figure 5.12 to find the maximum allowable group delay, $\tau_{\text{max}}$, between the amplitude and phase signal. Based on this number the minimum closed loop bandwidth, $f_{\text{CL,min}}$, must fulfill the following requirement.

   \[ \frac{1}{2\pi f_{\text{CL,min}}} < \tau_{\text{max}} \]

3) Trade off minimum PA gain, $G_{\text{PA}}$, for minimum filter bandwidth, $f_g$, using equation (5.31) repeated below:

   \[ f_{\text{CL}} \geq (1 + BG_{\text{PA},\text{min}} G_0)f_g \geq f_{\text{CL,min}} \]

Now that the minimum closed loop bandwidth, $f_{\text{CL,min}}$, is found we can according to equation (5.31) realize the necessary loop gain with by accepting a minimum PA gain. This limits the usable output amplitude interval see e.g. Figure 5.11 but is inevitable in a feedback loop were a minimum loop gain has to be ensured.

5.2.4 Gain compensation

In the previous section we showed that the bandwidth and loop-gain is a function of the PA-gain which again is a function of the amplitude of the modulated signal. For higher order systems, it’s not possible to chose the loop-gain arbitrarily high due to stability reasons. It is necessary to trade of bandwidth and loop-gain for gain stability margin.

To compensate for the varying gain in amplitude loop, we suggest to insert a variable gain stage after the calculation of the error, $e(t)$ (see Figure 5.13). Based on the PA-gain and the amplitude, the gain of the variable gain stage can be controlled to stabilize the loop-gain. The gain com-

![Figure 5.12: ACP versus amplitude delay](image-url)
compensation can be realized either as a two quadrant Gilbert multiplier or as an op-rational amplifier with the feedback resistance consisting of an array of switched resistors as shown Figure 5.14. In the former case a simple D/A converter would be required. In the latter case the signal \( g_c(t) \) would be a n-bit bus controlling the switches.

In both cases high precision of the gain value is not required as long as the loop-gain is kept high. This means that the linearisation system is insensitive to deviations of the PA-gain e.g. due to aging and temperature variations. The gain compensation can also be non-linear without affecting the system linearity. The gain control signal, \( g_c(t) \), could be based on a sample measurement of one PA because only an approximate value is required for the compensation. To ensure a low complexity of the D/A converter between the DSP and the multiplier the gain control signal should be quantized. The amplitude is mapped to intervals with an associated gain compensation value. The intervals and the gain control values should be chosen, so that the PA-gain times the gain control value has a minimum variation for all amplitudes. The actual gain values can be stored in a look-up table inside the DSP. The signal used for look-up is the amplitude which is already available in the DSP.

To demonstrate the principle, we simulated an amplitude loop with and without a gain control. We used the PA specified in Figure 5.11. The gain control values were quantized to 4 bits with a minimum value of 1 and maximum value of 5. The quantized gain control function and the resulting effective PA gain is shown in Figure 5.15. The maximum DC loop-gain was 20 and the filter pole was set to 1 MHz for both systems. The RMS amplitude error versus the maximum amplitude is shown in Figure 5.16. An improvement is observed for all amplitude levels even though both systems have the same maximum loop-gain and stability margin. For low amplitudes we observe an improvement on more than 6% RMS.
5.2.4.1 Simulation of the envelope FB system with gain control

The simulations carried out in previous section are described in the following. Since we wanted a precise simulation that incorporates the dynamics of the loop and the D/A converter for the loop gain control a transient simulation was carried out. The simulation procedure used is described in the section “Simulating the polar loop” on Page 67 with the modification that the loop has two inputs due to the gain control input (see Figure 5.17). This implies that an extra step has to be inserted in the simulation procedure described on Page 67. A step called (2b) should be inserted to calculate the gain control signal based on the input amplitude signal to the system. The way it was carried out here was that the gain control signal was calculated, based on the curve in Figure 5.15, and quantized to 4 bits. Then the signal was filtered with a RC lowpass filter. At this point both input signals to the envelope feedback system are found and the system can be simulated. The system was simulated in a math program [41]. The system Figure 5.17 was formulated as non-linear differential equation.

$$\frac{d}{dt} V_{pc}(t) = - \frac{1}{\tau} V_{pc}(t) - \frac{BG_0}{\tau} g_c[t, A(t)][V_{pc}(t)] + \frac{G_0}{\tau} g_c[t, A(t)] A(t) \quad (5.35)$$

The result of solving the differential equation was the power control signal to the PA. Based on that signal the output amplitude of the PA was calculated using the AM/AM (or rather the $V_{pc}$/...
AM) characteristic. The amplitude quality was evaluated as described in the previous section. A fragment of the transient simulation is showed to demonstrate what kind of information can be obtained. Figure 5.18 shows the amplitude (marked with red) and the PA gain. The compensated PA gain, which is the product of the PA gain and the compensation signal, is marked with circles. It can be seen that the PA gain is low for low amplitudes and for high amplitudes. When the PA gain is low we see that the compensated PA gain is increased. When the PA gain is high no further gain is injected into the loop. In this way no stability margin is violated.

5.2.5 Resume of polar modulation feedback

A new architecture for polar linearisation of RF power amplifiers suitable for digital transmitters was presented. The number of power consuming analog signal processing blocks were reduced and moved to the digital domain (DSP). Furthermore, an extension to the new architecture that reduces the stability constraints and maintains a constant loop-gain and system bandwidth was introduced. Simulations showed the reduction in the amplitude error was more 6% for low amplitude levels. The distortion sensitivity to delay difference in the signal path of the phase and amplitude was analysed, and simulations showing the relation between delay and ACP were shown. Design guidelines for the envelope feedback loop were given. Based on the fact that the polar loop can be treated as two separate feedback loops simulation guidelines which lower the simulation complexity were given. The simulations based on a radio system (GSM/EDGE) that supports

---

Figure 5.17: Envelope feedback system with gain control.

Figure 5.18: Envelope and gain transient simulation.
amplitude and phase modulation were presented. Signal quality (EVM) and spectral regrowth (ACP) were simulated based on a measured power amplifier. Equations that determine the settling of the Polar and Cartesian loop were found and the difference concerning the phase of the loop gain were discussed.

### 5.3 Impact of non-linearities on QPSK modulated signal

In the following the influence of different amplitude and phase non-linearities of the transmitter stage is investigated. The signal quality is measured in terms of ACP and RMS vector error. First, simple idealised non-linearities are applied to the transmitter stage to show the amount of non-linearity necessary to cause a certain spectral re-growth and in-band vector error. Then the characteristic of a simulated CMOS power amplifier is applied to demonstrate the performance of a PA not optimised to linear operation.

The signal quality received depends both on the type of modulation and the non-linearity in the transmitter stage. The modulation used in the following is a QPSK modulation with a root raised cosine puls-shape filter employed on both the sender and receiver side. The symbol rate is 4.096 Mchip/s. The puls-shape factor is 0.22.

The AM/AM and AM/PM characteristics used to model the PAs are shown in Figure 5.20 and Figure 5.21. The red curves are the characteristics of a simulated 1W CMOS power amplifier.

The amplifier is non-linear in phase as well as amplitude. An linear amplifier would have a straight line as AM/AM characteristic like the blue graph in Figure 5.20 and a horizontal AM/PM characteristic similar to the black line in Figure 5.21. To get a feeling how much non-line-
arity is necessary to cause a certain distortion of the modulated signal two other PA models are simulated. One model contain a phase non-linearity and no amplitude non-linearity. The other contains a amplitude non-linearity only.

### 5.3.1 Phase non-linearity

In the first simulation a phase error is introduced. The phase error is proportional to the input amplitude (see blue graph in Figure 5.21). The AM/PM characteristic is a straight line. The maximum phase error was swept and the resulting ACP error is depicted in Figure 5.22. The requirements of the signal quality is taken from the early specification of wide band CDMA specification [42]. The required ACP must be better than 40dBc and the RMS vector error better than 12%. It is seen that the ACP is better than 40dB for a maximum phase error of less than 8 degrees. If the ACP should be below 50 dB the maximum phase error should be below 2.5 degrees. This is hard to fulfil without using power-back-off on the power amplifier or without using linearisation. The magnitude vector error is plotted (Figure 5.23) and we see that the EVM requirement is critical to a phase error than the ACP requirement. The distortion limit of 12% is reached at 110 degrees after the ACP limit. So it seems to be the ACP is the critically parameter and it the determines the maximum allowable phase error for this system. The error vector is based on the received signal and a undistorted reference signal each normalised to the unit power as explained Chapter 1.3.5.
5.3.2 Amplitude non-linearity

In the next experiment we model the system with at third order non-linearity (see Figure 5.24). The model has one free parameter determining the third order non-linearity. The function is scaled to limit the maximum value to 1.

The amplitude function is as follows

\[ AMAM(x) = x - ax^3 \]  

The function has a zero at \( x = 0 \) and \( x = 1/(\sqrt[3]{a}) \). The toppoint is at \( (x, y) = (1/(\sqrt[3]{3a}), 2/(3\sqrt[3]{3a})) \). The function will only be used for \( a \in \{0;1/2\} \) as the curve approximates a AM/AM characteristic reasonable in this interval. The function is shown for three different parameters in Figure 5.24. In Figure 5.25 it is seen that an ACP limit of 40dB is reached at \( a=0.13 \), and 50dB limit is reached at \( a=0.04 \).

The mean vector error stays below 7% in the simulated interval which is below the system specification of 12%. The simulation again shows that the ACP requirement is the most critically.

The input signal of the non-linearity had the following properties. A maximum amplitude of 1, an average power of 0.195W (in 1 Ohm) and a maximum power of 0.5W. The peak to average power is \( 2.56 = 4.1 \text{dBW} \).
5.3.3 Distortion of a simulated CMOS PA

The distortion based on a simulated 2GHz, 0.25u CMOS amplifier is calculated. The AM/AM and AM/PM function are showed in Figure 5.20 and Figure 5.21 respectively. The output power is 1W but the AM/AM characteristic is normalised to one. The AM/AM characteristic is close to linear until amplitudes of 0.2 where it starts to compress. The AM/PM characteristic has a huge phase change for small amplitudes and settles around 100 degrees after 0.1 Volt. Remember that a fixed phase corresponds to a fixed delay and causes no distortion. So the amplifier is dominated by phase distortion until around 0.15 Volt where the amplitude distortion gets domination. It is observed that a local optimum in the ACP and error vector magnitude at 0.2 Volt where the phase distortion has ceased and the amplitude distortion is still weak.

It is amazing that the vector error is small if one consider the tractory in the IQ-plane (Figure 5.29). The PA output (red curve) is close to the unit circle most of the time. This is due to the large small signal gain and the compression. Most signal amplitudes will be amplified and compressed so the final amplitude is close to the maximum output of one volt. The reason that one can achieve a relative low error vector magnitude is due to the receive filter (root raised cosine filter). It compensate for a lot of distortion as long as the tractory gets near the constellation points. The filter removes some of the excursion so the tractory becomes similar to the non-distorted tractory. Compare for instance the un-filtered (Figure 5.29) and the filtered (Figure 5.30) PA signal. The advantage of the receive filter was also confirmed by simulations where the non-linearity limited the minimum amplitude. It was there not possible to pass the origin in the trac-
tory plot. After applying the receive filter the tractory will pass through the origin and the tractory will be similar to the undistorted reference signal with a small error vector magnitude as result.

The aligned signal and the reference signal is depicted in Figure 5.30. The reference signal is the PA input signal filtered with the receive filter. So at this point the transmit and receive filtering corresponds to a Nyquist filter and the IQ-constellation points are easily recognize. The receiver adjust the phase based or the complex correlation between the received signal and a reference signal. After that, the power of the signal is normalised to 1 as in the case of the reference signal. This procedure ensures a low vector error (see Chapter 1.3.5).

5.3.4 Summary of impact of non-linearities on QPSK modulated signal

Three simulations were carried out. One with a phase non-linearity only, one with a amplitude non-linearity only and a simulation with phase and amplitude non-linearity. It was seen that the phase non-linearity as well as amplitude non-linearities causes severe degradation in ACP (adjacent channel power) separately. It was also observed that the ACP specification is the most limiting requirement to the signal quality compared to the error vector magnitude. The evaluation was based on a simple system specification of ACP>40dB and a mean vector error of 12%.

Figure 5.27: ACP, amplitude distortion versus input voltage

Figure 5.28: Relative magnitude error in percent versus input amplitude
5.3.5 Remarks to the simulations

The simulations were carried out in MathCad [41]. 128 symbols were simulated with a symbol oversampling of 32. Random information bits for the inphase and quadrature channel were generated and the oversampled signal was interpolated using a root raised cosine filter with alpha equal to 0.22. The symbol rate was 4.096 Mchip/s and the sampling frequency was 131MHz. The complex baseband signal was normalised to a maximum amplitude of 1 volt and applied to
the non-linearities presented in the previous chapters. The receive filter was a root raised cosine filter. The ACP calculations were based on the power in one neighbour channel and the power of the wanted channel. The calculation of the vector error signal is based on the received signal and a reference signal. The reference signal has been exposed for the same filtering as the received signal but has not experienced any non-linearity. The two signals was normalised to 1 Watt (in one Ohm) before there where subtracted.
Chapter 6
Design of key building blocks

In this chapter some of the critical components in a linearisation system will be analysed, designed and measured. The work will focus on polar modulation feed back loop. First different types of published amplitude detectors are presented (Chapter 6.1) then some on-chip realisations of power amplifiers with AM-detection will be discussed (Chapter 6.2). In Chapter 6.6 experiments with commercial RF components for the polar linearisation loop are carried out to gain some practical experience and recognize the design challenges. These experiments deal with linear amplitude detection and phase shifters which are relevant not only for the new architectures presented in Chapter 5, but also for the traditional polar loop transmitter. In Chapter 6.9 a CMOS power amplifier design is described from design to measurement. The last chapters treat three different amplitude detectors. Two of the non-linear detectors have been fabricated and measured whereas the linear detector has been delayed from fabrication so measurements on this chip were not possible.

The work documented in Chapter 6 has resulted in a number of publications which can be found in [43][44][45][46][47][48]. Some of them can be found in Appendix D.

6.1 Introduction to AM-detectors

There are several ways to realize an amplitude detector. In the following the principle of amplitude detection is presented.

6.1.1 Diode AM-detectors

The most common way to realize an amplitude detector is based on the diode as the non-linear element. The strong non-linear relationship between the current and voltage is shown in Figure 6.1 and equation (6.1) below.

\[ I_{diode} = I_s \left[ e^{\frac{V_{diode}}{V_n}} - 1 \right] \]  

(6.1)

This characteristic is often approximated with an ideal voltage controlled switch that is closed only for positive voltages. Although the current is negative for negative voltages, the current is
very small and can in most applications be neglected. The temperature dependencies are modelled through the saturation current, $I_s$, and the thermal voltage, $V_t$.

$$V_t = \frac{kT}{q} \quad (6.2)$$

$$I_s = I_s(T_0) \left(\frac{T}{T_0}\right)^{\frac{2}{n}} \exp\left[-\frac{q\psi}{k} \left(\frac{1}{T} - \frac{1}{T_0}\right)\right] \quad (6.3)$$

Perceiving the diode as a switch is convenient when some circuits have to be understood. Among them are the rectifying circuits shown in Figure 6.2. The circuit in Figure 6.2(c) is a voltage multiplier which means it detects twice the envelope value.

The diode in Figure 6.2(a) and Figure 6.2(d) conducts currents for positive input voltages. This means that the capacitor is charged in each positive half period of the carrier until the capacitor voltage is equal to the envelope value. The resistor is discharging the capacitor so the envelope can be tracked for decreasing envelope. The RC time constant must be $n$ times longer than the
period of the carrier to ensure a stable voltage between adjacent peaks of the RF carrier. Further
it must be ensured that the envelope signal is not attenuated by the lowpass filter function of the
RC components. A design rule for the maximum time constant used in AM detection, based on
the modulation index and the maximum modulation frequency is found in [49]. The AM-modu-
ulated carrier is described in the following equation.

\[ s(t) = [1 - (m \cdot A(t))] \sin(\omega t) \quad \text{and} \quad 0 \leq A(t) \leq 1 \quad (6.4) \]

where

- \( A(t) \) is the modulating signal band limited to \( f_{\text{max}} \)
- \( m \) is the modulation index, \( 0 \leq m \leq 1 \)

The time constant of the AM-detector is then restricted to the following.

\[ RC < \frac{m^2 - 1}{3.8 f_{\text{max}}} \quad (6.5) \]

In general it is difficult to find a closed loop expression for the behaviour of the AM-detectors,
but with some restrictions it is possible to calculate expressions that explain many of the phe-
nomena observed in a practical AM diode detector. Consider for instance the circuit
Figure 6.2(d). If the circuit is connected to a generator with a non-zero output impedance as
shown in Figure 6.3 it is possible to calculate the corresponding DC voltage[50].

![Figure 6.3: AM detection with biased diode.](image)

The calculations are based on equation (6.1) and the Ritz-Galékin method [51]. The relation
between the detected voltage and the input power is given below.

\[ Bessel_0 \left( \frac{R_g P_{\text{inc}}}{n V_t^2} \right) = \left( 1 + \frac{I_0}{I_s} + \frac{V_{\text{out}}}{R_L I_s} \right) e^{\left( 1 + \frac{R_g + R_L}{R_L} \right) \frac{V_{\text{out}}}{n V_t} + \frac{R_L I_0}{n V_t}} \quad (6.6) \]

where

- \( P_{\text{inc}} \) is the power delivered in a load matched to the generator \( V_g^2 / (8 R_g) \).
- \( n \) is emission coefficient of the diode (see equation (6.1))
- \( I_s \) is the saturation current in equation (6.1) and (6.3)
- \( I_0 \) is the diode bias-current
- \( R_L \) is the load resistance

This equation can be used to show some of the important relationships between output voltage,
load resistance, bias-current and temperature. The characteristic of the detector is shown in
Figure 6.4. The traditional characteristic for high resistive load follows the square law for small
signals and the linear law for high signals. The figure also shows curves for finite and small load
resistances where a higher order transition exist between the square law region and the linear region. The deviation from the traditional characteristic can be minimized by applying bias current.

Figure 6.4: Transfer characteristic of diode detector for different loads.

Figure 6.5: Published FET AM-detectors. (a) and (b) MOSFET MMIC AM-detectors (c) GaAs AM-detector.
6.1.2 Field Effect Transistor AM-detectors

The traditional way to design an AM-detector is to use a diode. However, it is possible to use field effect transistors as a non-linear element. A higher sensitivity can be achieved for GaAs-FETs ([52]) and MOSFETs ([53]) detectors compared to a Schottky diode detector. But also the noise performance and temperature stability is better [52], [53]. 1/f noise is present for higher power levels but is insignificant. For low detection levels the dynamic is set by white noise[52]. The topologies used are shown in Figure 6.5. The FET offers an extra possibility to increase the sensitivity or improve matching by biasing the gate. It is also possible to apply a low frequency sinusoid at the gate. The output is in this case amplitude modulated by the low frequency signal. The advantage is according to [52] that offset can be avoided.

The technology in [53] was an MMIC technology with a low integration level.

6.1.3 Linear AM-detectors

Linear detectors can be realized in several ways. One method is to use a limiting amplifier and a switching mixer. The principle of the detector is shown in Figure 6.6. The input signal is split into two signals. One is led through a limiter to obtain a constant envelope signal for the LO input of the mixer. The constant envelope signal and the original signal are multiplied and low-pass filtered. The result is the envelope of the RF signal.

Another method is to use two identical non-linear elements (Figure 6.7). One is used to detect the envelope of the RF signal. The other is used to build the inverse function. This can be carried out using an operational amplifier to realise feed back around the second non-linear element. This principle was used in a CMOS design [54] for envelope elimination and restoration.
6.2 Architectures and design considerations for envelope feed back

It is an advantage to integrate as much functionality as possible on a chip to reduce costs when designing a polar linearisation system. Depending on the technology for the PA two strategies for maximum integration are possible:

1) A two chip solution with the PA as one chip realised in a special technology optimized for power efficiency and linearity. A second chip is then used to integrate all parts of the linearisation system. A suitable technology for the second chip would be a CMOS or BIC-MOS technology.

2) A one chip solution where the PA and linearisation system are integrated on the same chip. CMOS or BICMOS technology would be a good choice because analog signal processing can be performed and analog options such as voltage independent resistors and thin oxide capacitors can be made available.

Depending on the chosen solution different methods to sense the envelope of the PA (see Figure 6.8) must be used. Either a coupler is used between the matching circuit and the antenna or the signal is sensed directly on-chip using for instance a resistor as shown in Figure 6.8. The impedance level at this place is low (1-5\,\Omega). In order not to tap too much power and thereby degrade the power efficiency of the output stage the sense resistor should be chosen significantly larger than the load resistance of the PA (R_{load}). In the figure we have chosen a ratio of 10. This results in an acceptable degradation of the power efficiency assuming that voltage swing is unchanged by the resistor, R_{tap}. The new efficiency after tapping power can be calculated as follows.

\[
\eta_{tapped} = \frac{1}{\eta + \frac{P_{tap}}{P_{out, untapped}}} = \frac{1}{\eta + \frac{R_{load}}{R_{tap}}} = \frac{1}{\eta + \frac{R_{tap}}{R_{load}}} \quad (6.7)
\]

If an output stage has an efficiency of 50\% then the resulting efficiency with a 10 times larger sense resistance, R_{tap}, degrades the efficiency to 47.6\%. In a linearisation system other circuits have to be taken into account which will degrade the overall efficiency even more. This may seem drastic, but it should be compared to the case where the loss in efficiency is caused by
backing the PA off with 3 to 6 dB. This often causes a degradation of 20-30% or more. Tapping or sensing directly on the output transistor implies that many harmonics may be present. This is the case for class B,C,E power amplifiers. Especially as the amplifier approaches the compression point the harmonic content will increase. This would lead to an erroneous detection of the amplitude of the fundamental frequency. Therefore a bandpass filter should be applied to select the fundamental frequency. Also the amplitude is lower at the output-transistor due to the impedance level of the PA. The amplitude is scaled by the square root of the impedance levels. So if the wanted load resistance of the PA is 2 Ohm the amplitude will be 20% of the amplitude at a nominal 50 Ohm load.

\[ A_{PA} = \sqrt{\frac{R_{PA}}{R_{50}}} A_{R_{50}} = \sqrt{\frac{2}{50}} A_{R_{50}} = 20\% A_{R_{50}} \quad (6.8) \]

### 6.2.1 The Bandpass filter

The bandpass filter in Figure 6.8 must be a simple realisation to avoid too many losses at high frequencies (1-2GHz). The obvious choice consists of a DC separation with a capacitor, \( C_1 \), and a LC resonator tank (\( L, C \)) selecting the carrier frequency. If we use an on-chip inductor the maximum quality factor of the resonator will be approximately 5. This loss will be the dominant loss and is represented by \( R_p \) in Figure 6.9. The value can be expressed by the inductor quality, \( Q_L \).

\[ R_p = R_L (Q_L + 1) \quad \text{and} \quad R_L = \frac{\omega_0 L}{Q_L} \Rightarrow \]

\[ R_p = \omega_0 L \left( Q_L + \frac{1}{Q_L} \right) = \omega_0 L Q_L \quad (6.10) \]

To find a typical value for \( R_p \) let’s assume a quality factor of 5 at 2GHz and an inductor value of 1nH. The equivalent parallel resistance is 62Ω. The purpose of resistor \( R_f \) in Figure 6.9 is to protect against draining too much power from the PA and to maintain the quality factor of the tank. If the loss resistor, \( R_{fp} \), is high enough to prevent degradation in power efficiency (refer to equation (6.7)) the resistor \( R_f \) can be avoided. This would also give us the advantage of a better cancellation of harmonic currents from the output transistor. If a high resistance is necessary for \( R_f \) the tank will only filter out the harmonics for the amplitude detector.
In order to minimize the losses in the tank the product $LQ_L$ must be maximized (see equation (6.10)). The quality factor will not limit the bandwidth because $Q_L$ is limited by the CMOS technology to 5-6 which yield a bandwidth of approximately 400MHz at 2GHz. A quality factor of 5 leads to a lower tank impedance of a factor 7.6 and 13.4 of the 2nd and 3rd harmonic respectively. Thereby the influence of the harmonics is reduced and the fundamental frequency component is detected. The reduction of the tank impedance can be derived by calculating the tank impedance at the harmonic frequencies and dividing by the impedance at resonance. The tank impedance is

$$R_p \frac{(j \cdot \frac{\omega}{f_0}) \frac{1}{Q}}{(j \cdot \frac{\omega}{f_0})^2 + (j \cdot \frac{\omega}{f_0}) \frac{1}{Q} + 1}$$

(6.11)

where the quality factor and the resonance frequency are given by

$$Q = \frac{R_p}{2\pi f_0 L} = \frac{1}{2\pi \sqrt{LC}}$$

$$f_0 = \frac{1}{2\pi \sqrt{LC}}$$

(6.12)

The impedance of the tank is then

$$|Z_t(f, f_0)| = R_p \frac{1}{\sqrt{Q^2 \left(\frac{f}{f_0} - \frac{1}{f}\right)^2 + 1}}$$

(6.13)

At the resonance, $f_0$, this simplifies to the parallel resistor of the tank, $R_p$. At the harmonics of the resonance frequency the impedance is reduced by the following factor.

$$\frac{|Z_t(f_0, f_0)|}{|Z_t(n \cdot f_0, f_0)|} = \sqrt{Q^2 \left(n - \frac{1}{n}\right)^2 + 1}$$

(6.14)

The tank impedance reduction factor can be found in Table 6.1 for different harmonics and quality factors.

<table>
<thead>
<tr>
<th>Harmonic</th>
<th>Q=3</th>
<th>Q=4</th>
<th>Q=5</th>
<th>Q=6</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=2</td>
<td>4.6</td>
<td>6.1</td>
<td>7.6</td>
<td>9.1</td>
</tr>
<tr>
<td>n=3</td>
<td>8.1</td>
<td>10.7</td>
<td>13.4</td>
<td>16</td>
</tr>
<tr>
<td>n=4</td>
<td>11.3</td>
<td>15</td>
<td>18.8</td>
<td>22.5</td>
</tr>
<tr>
<td>n=5</td>
<td>14.4</td>
<td>19.2</td>
<td>24</td>
<td>28.8</td>
</tr>
</tbody>
</table>

6.3 Envelope feedback system with non-linear AM-detectors

Most of the previous mentioned AM-detectors are non-linear. So to ensure a high linearity of the system we have to rely on high gain feedback and matching of detectors. Consider the sys-
tem of Figure 6.10. The characteristic of the AM-detector is represented by the function $g_2$. We can calculate the error signal $e(t)$ and assume it is zero due to the fact that the loop gain is high\(^1\).

$$e(t) = A_{ref}(t) - g_2(A_{PA}(t)\alpha) \approx 0$$  \hspace{1cm} (6.15)

We can solve the equation for the amplitude of the PA using the inverse function of $g_2(x)$.

$$A_{PA}(t) = g_2^{-1}(A_{ref}(t)) \cdot \frac{1}{\alpha}$$  \hspace{1cm} (6.16)

If the function $g_2$ is linear the inverse function $g_2^{-1}$ will be linear and the system will behave linearly. But if $g_2$ is non-linear the whole system will behave non-linearly due to the fact that $g_2^{-1}$ is non-linear. We need to compensate for the non-linearity $g_2^{-1}$ in equation (6.16). This can be carried out by generating an RF reference signal, $A_{ref}(t)$, and detect it using a similar AM-detector with the same characteristic $g_2(x)$ as shown in Figure 6.11.

$$A_{\alpha}(t) = g_2(\beta \cdot A_{ref}(t))$$  \hspace{1cm} (6.17)

$\beta$ represents the gain of an amplifier that can be necessary if the dynamic range is limited. Replacing $A_{ref}(t)$ equation in (6.16) with $A_{\alpha}(t)$ in (6.17) we obtain a linear equation and thereby a linear system. So the main point when using non-linear AM-detectors is to use the matching between the detectors.

$$A_{PA}(t) = g_2^{-1}(g_2(\beta \cdot A_{ref}(t))) \cdot \frac{1}{\alpha} = A_{ref}(t) \cdot \frac{\beta}{\alpha}$$  \hspace{1cm} (6.18)

Another important observation is that the gain of the system is set by the constant factor $\alpha/\beta$. The matching requires the same input level to the AM-detectors. Notice at the amplifier and the attenuator in Figure 6.11 must be linear unless they are identical in which they can be incorporated as a part of the non-linear detector.

---

1. The error is only zero if an integration exists inside the feedback loop and the amplitude signal is constant. Since the error is small due to a high loop gain the error caused by this assumption is negligible.
6.4 Experiments with discrete RF components

This section deals with some practical experiments and measurements on functional blocks to a linearisation system. The purpose of the work has been to gain experience of RF design and learn the behaviour of commercial RF components. The experiments to be performed will reveal some advantages and drawbacks of the building blocks. Discrete commercial available parts are used which implies that some principles and problems are treated differently than it would be in the case of a fully custom integrated chip design. However, the experiments are expected to reveal fundamental properties and challenges of RF-systems and feed back systems.

6.4.1 The system

The building blocks are intended to be used in a Polar Modulation Feed Back system such as the one proposed in Chapter 5.2.1. The architecture with a phase shifter is shown in Figure 6.12. The idea of using a phase shifter was tested and proved to be efficient to reduce ACP in [9]. The envelope of the PA output is extracted and compared with a reference signal of generated by the DSP. The error is amplified and fed into the power control input of the power amplifier. The feed back ensures that the PA output has the right amplitude.
The phase feed-back loop consist of a phase-shifter in series with the PA, and a phase detector that extracts the phase error between the PA output and a reference signal. The output of the phase shifter is the phase error. This error is used to control the phase shifter. If the loop gain is high, the phase of the PA-output signal will be equal to the phase of the reference signal plus a fixed offset.

The system in Figure 6.12 has to be realized with commercial available components. This leads to a slightly different system with more components compared to an on-chip realisation. Some of the functionality can only be realized by combining several components. We need to split the modulated RF signal into a RF constant envelope signal and a low frequency envelope signal. The constant envelope signal is used as a reference signal for the phase detector (see Figure 6.12). The phase detector is realized as a mixer so no amplitude variations can be tolerated. The generation of the necessary phase modulated RF signal and envelope signal used as input for the linearisation system can be derived from a fully modulated RF signal coming from a commercial RF generator as shown in Figure 6.13.

The system in Figure 6.12 is realized by standard RF components and two printed circuits boards specially made for the system as shown in Figure 6.14. The limiter is used both for amplitude detector and for the phase detector. The phase detector consist of a limiter, a mixer and a low pass filter. When two constant envelope signals are mixed the lowpass component of the output is a function of the phase difference. The amplitude detector is realized as a mixer and a limiter. The limiter generates a constant envelope of the PA-output signal that is mixed with the PA-output. The low frequency product of the mixer represents the magnitude of the RF envelope. This kind of detector is in theory linear. Two PCB boards must be designed to provide gain and add an offset.

Figure 6.13: Generation of the reference signals.

Figure 6.14: Linearisation system with commercial available parts.
The system consists mainly of Mini Circuits components with SMA connectors. The RF inputs of the PCB are also equipped with SMA connectors. All parts are connected by coaxial cables. The power amplifier (RFMD2108) is a CDMA amplifier evaluation board from RFMD. Two of the circuit boards handle the extraction and amplification of the error in the envelope and the phase loop. The two boards are similar in the sense that they both get their input from a mixer, and they both amplify an error. The board for the phase loop does not have an input because the phase difference should be held constant to eliminate the phase distortion. The phase of the PA should just follow the phase of the reference signal. The board for the power loop has an input so the magnitude can be controlled (see also Figure 6.13).

6.4.2 Characterisation of a commercial PA

In the following measurements of the PA are presented and explained. The PA chosen for the linearisation system is a RFMD 2108 from RF Micro Device. This PA is designed for the American CDMA and AMPS system and other applications in the band from 800MHz to 950MHz. The amplifier is designated as a linear amplifier and therefore supports amplitude modulation. The immediate choice for the linearisation system would be a non-linear amplifier because of the power efficiency. This type of amplifier was not available so a linear was chosen. The PA can however act as a non-linear PA either by over-driving the input signal or by over-driving the bias control voltage ($V_{pc}$). Further, to be linear in the contents of a CDMA system does not imply linear in mathematical sense. So the purpose of the linearisation system is to improve the linearity of the given PA. In Table 6.2 some of the PA specifications are shown.

<table>
<thead>
<tr>
<th>Table 6.2: Specification of the PA RFMD 2108</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Feature</strong></td>
</tr>
<tr>
<td>Supply</td>
</tr>
<tr>
<td>Output Power CW</td>
</tr>
<tr>
<td>Linear Output power (CDMA)</td>
</tr>
<tr>
<td>Efficiency</td>
</tr>
<tr>
<td>Frequency</td>
</tr>
<tr>
<td>ACP 885 kHz offset</td>
</tr>
<tr>
<td>ACP 1.98 MHz offset</td>
</tr>
</tbody>
</table>

Notice that the ACP is measured in offsets of 885kHz and 1980kHz. In the following these measurements are made for WB-CDMA signals with a channel spacing of 5MHz, as this is the system focused on.

In the following two parameters are swept. First the input power is swept while the power control voltage is kept constant. The adjacent channel power and the gain are measured. Then the power control voltage is swept while the input power is held constant.

6.4.2.1 Output versus input power

Measurements of the output power and the ACP was made by sweeping the input power of the PA as shown in Figure 6.15. The input signal was a WB-CDMA signal (4096 MChips/s, $\alpha = 0.22$) with a channel spacing of 5MHz generated by R&S SMIQ03. The power control
The signal ($V_{pc}$) was set to 3.6 Volt. The supply voltage was 4.8 Volt. The input power was swept from -5dBm to 3dBm. The output power is shown Figure 6.16. The PA compresses slightly near the maximum output power level (31.5dBm). This level is above the specified maximum linear output power so the control voltage ($V_{pc}$) has to be lowered to achieve the specified linearity. The small signal gain is 30.2dB. The gain decreases from 30.3dB down to 16.7dB. This variation of the gain will reduce the achievable bandwidth of the linearisation system due to stability requirements. A measurement of the adjacent channel protection (ACP) is shown in Figure 6.17. A maximum is achieved at an input power level of -4dBm. This maximum is also a function of the power control ($V_{pc}$) as this influences the operation point of the transistors in the PA. The PA was specified to -47dB ACP which holds until the input power exceeds -3dBm.
This corresponds to 26.5 dBm output power (see Figure 6.16). The specified linear output power is 28 dBm so according to the specification the PA should be able to deliver 1.5 dB more output power with an ACP of 47 dB. A reason for the deterioration could be that the measurements were based on 5 MHz bandwidth and $\alpha=0.22$ which differs from the assumption that underlies Table 6.2.

### 6.4.2.2 Output power controlled by the power control signal

In the following measurement the input power was set to -4 dBm and the power control of the PA was swept. The supply voltage was 4.8 V and input signal was a WB-CDMA signal as described in previous Section 6.4.2.1. Figure 6.18 shows the result of the sweep. The maximum output corresponds to an output power of 25.8 dBm. The gain\(^1\) in Figure 6.19 ranges from 7.4 to 0.6 which corresponds to approximately 21.6 dB. The big gain variation is due to the fact that

![Figure 6.18: Output power versus power control voltage.](image1)

![Figure 6.19: Gain versus control voltage.](image2)

the control voltage, $V_{pc}$, determines the bias point of the PA. The big gain variation makes it difficult to stabilize a linearisation system. At high output power the gain is low which lowers the loop gain of a linearisation system. Therefore the relative error will be higher at high output levels. However the ACP of the PA itself increases with increasing output power (see

\(^1\)The gain was calculated based on a average power measurement of a WB-CDMA signal. This is not according to a normal definition based on a sinusoidal but it still gives a indication of the gain as long as the peak to average ratio is low.
which in this case can be perceived as a compensating effect. In general one can say that it is a desirable feature that the ACP increases (better quality signal) as the output power increases, because the total power in the adjacent channel will tend to be low for all output levels.

6.4.2.3 Summary of PA measurements

In this section the output power of the PA was measured sweeping the input power and the control voltage \( V_{pc} \) independently. The small signal gain and the ACP was measured. A large small signal gain variation of 21.6dB was observed when controlling the power through the control voltage \( V_{pc} \). The large variation will reduce the achievable bandwidth of a linearisation system due to stability requirements. Therefore much attention should be paid to the power control when designing a PA. The easiest way to linearise using predistortion would be to predistort the power control input signal compared to the case where the RF input signal of the PA is predistorted. This is clear due to the fact that the power control signal is a lowpass signal (base band).

6.4.3 Amplitude detection

In this chapter different variations of a amplitude detector are measured. The principle of the detector planned for the linearisation system is shown in Figure 6.6 and repeated below.

![Figure 6.21: Principle of amplitude detection.](image)

The input signal is split into two signals. One is led through a limiter to obtain a constant envelope signal. The constant envelope signal and the original signal are multiplied and low-pass filtered. The result is the envelope of the RF signal.

Three different realisations of the amplitude detector was measured to find an adequate linear detector. The first problem encountered was the limiter from Mini Circuits (PLS-2) which had a surprisingly poor clipping ability. A sweep of the input power is shown in Figure 6.22. This
non-flat characteristic gave reason to believe that it would not be possible to show an improvement with this limiter in the linearisation system. Still the intended detector was measured and the result is shown in Figure 6.23. It is obvious that the detector is non-linear. Note that the curve should only be considered for input values greater than 0.35 as the limiter is specified only for a 12 dB input range (3-15 dBm). The first attempt to improve the characteristic was to insert an amplifier after the limiter (see Figure 6.24) to get a stronger constant envelope signal. The data sheet for the used mixer (Mini Circuits ZEM-4300) was specified for power levels of 7-
10dBm. Since the limiter only delivers -5dBm one could hope for improvement by using the same power level as used for the specification in the data sheet. Figure 6.25 shows the result. The characteristic is still not linear although it could be useable for input values between 0.2 volt to 0.4 volt. The limiter will always influence the linearity as long as the mixer behaves as a multiplication.

This also led to the last experiment (Figure 6.26) where the limiter was removed. If it is assumed that the mixer works as a switching mixer, the mixer output will be independent of the actual LO level as long as it is high enough. This also implies that the output level of the mixer depends on the RF input of the mixer only. So the DC output of the mixer will in theory depend linearly of the amplitude of the input signal to the amplitude detector.

It is seen that this yields the highest linearity (Figure 6.27). So this amplitude detector will be used for the linearisation system. For the comparison the three characteristics for the different amplitude detectors are shown below. All curves are normalized.

**6.4.4 Circuit boards**

To complete the linearisation system four circuit boards were designed. The linearisation system consists of two feedback loops (see Figure 6.14). In both the amplitude and the phase loop a circuit board is needed to subtract two signals and amplify the error. These two boards process base band signals. Two RF circuit boards were designed. One with a RF phase shifter and one RF gain stage for general purposes.
6.4.5 Error amplifier for the feedback loops

The subtraction and amplification in the power loop were performed with a circuit, with the functionality shown in Figure 6.29. The input signals consist of the wanted signal, \( A \), and the output from the amplitude detector, \( A_{PA} \), (see also Figure 6.14 on Page 91). Both inputs are terminated with 50Ω near the SMA connector to avoid reflections. \( R_1 \) and \( R_2 \) are chosen significantly larger than the 50Ω to ensure proper termination. Further the input from the amplitude detector has an inductor in series to suppress the upper mixing product from the amplitude detector.

The transfer function of the first gain stage is shown below.

\[
V_1 = (1 + g_2)\alpha \cdot A_{PA} - g_2 A
\]  

(6.19)
where

\[ g_2 \] is the gain for the negative input, \( g_2 = \frac{R_3}{R_2} \).

\[ \alpha \] is the attenuation from the positive input to the positive terminal of the operational amplifier.

\[ \alpha = \frac{R_{P1}}{sL + R_1 + R_{P1}} \]  \hspace{1cm} (6.20)

In normal operation the loop gain is high (>10) and one can assume that the loop error \( V_1 \) is equal to zero. Under this condition one can express the output of the amplitude detector as a function of the amplitude control signal \( A \).

\[ A_{PA} = \frac{g_2}{(1 + g_2)\alpha} \cdot A \]  \hspace{1cm} (6.21)

The output amplitude of the PA is controlled by \( A \) because \( A_{PA} \) is proportional to the amplitude of the PA output.

The PCB was built and successfully tested for the functionality described above. The component values are shown in Table 6.3. The circuit board for the phase loop is very similar to the one for the power loop. The only difference is how the signal \( A_{PA} \) is scaled using the potentiometer (Figure 6.30).

**Table 6.3: Component values for power loop circuit board.**

<table>
<thead>
<tr>
<th>R1 5kΩ</th>
<th>R2 5kΩ</th>
<th>R3 5kΩ</th>
</tr>
</thead>
<tbody>
<tr>
<td>R4 2.5kΩ</td>
<td>R5 2.5kΩ</td>
<td>R6 5kΩ</td>
</tr>
<tr>
<td>P1 200kΩ</td>
<td>P2 200kΩ</td>
<td>P3 50kΩ</td>
</tr>
<tr>
<td>P4 50kΩ</td>
<td>OP1 LM6164</td>
<td>OP2 LM6164</td>
</tr>
<tr>
<td>C1 1nF</td>
<td>C2 10 uF</td>
<td>C3 1uF</td>
</tr>
<tr>
<td>C4 10 uF</td>
<td>C5 1uF</td>
<td>L1 180 uH</td>
</tr>
</tbody>
</table>
6.4.6 Phase shifter

The phase shifter is based on a parallel LC resonator where the capacitance is controlled by a varactor diode (see Figure 6.31). The actual phase of the RF signal can be changed by changing the resonance frequency of the resonator. The amplitude of the signal will also change but if the signal is clipped using a cascade of amplifiers a constant amplitude will be obtained. The phase shifter was realized with a buffer amplifier at the input and at the output to decouple the resonator from adjacent stages. Further amplifiers can be connected externally to perform the clipping. The resonator was further decoupled by 1 kΩ resistors to ensure a high Q of the tank. The tank itself was simulated in Aplac[55] (see Figure 6.32). Package models were used for the inductor, capacitor and the diode. It is seen that the phase varies at least 90 degrees as the bias voltage of the varactor is swept from 2 V to 4.5 V. The magnitudes and phases are simulated for three frequencies, 870 MHz, 890 MHz and 910 MHz. It is seen that the phase shift is obtained for all frequencies. The tank amplitude changes at most 6 dB. It is seen that the attenuation at the output varies between 38 dB and 50 dB from input to output. The loss from the input buffer (gain=1 in simulation) to the tank varies between 8 dB and 12 dB. The largest loss contribution originates from the voltage divider which divides by 20 (26 dB). It is possible to reduce the loss significantly by inserting a matching circuit that transforms the impedance down and preserves the power. How-

![Figure 6.30: Circuit board for subtraction and amplification in the phase loop.](image)

![Figure 6.31: Schematic of phase shifter.](image)
ever, this was not done in this realisation because the loss will be compensated by an extra gain stage. The photo of the phase shifter is shown in Figure 6.33.

6.4.6.1 Measurement of the phase shifter

The phase shifter was tested as shown in Figure 6.34. Two Rohde & Schwarz RF generators were synchronised using a coaxial cable. Any phase change in the signal path from generator 1 to the mixer will be detected as a voltage change on $V_a$. To ensure that amplitude-changes originating from the phase shifter will not influence the phase measurement a limiter was inserted. As we experienced, the limiter was not clipping perfectly (see Section 6.4.3 "Amplitude detection" on Page 95) so any minor voltage change was measured with the analyser ($V_b$). These changes can be used to correct the phase measurement. The measurement procedure is as follows. First the phase between the generators are swept to verify the functionality of the phase detector. Then the phase of the generators is fixed and the control voltage of the phase shifter ($V_c$) is swept. During the sweeps the voltages $V_a$ and $V_b$ are measured. In Figure 6.35 the result of sweeping the phase difference of the generators is shown. The output is as expected a part of a sine function. This will always be the result when two sinusoidal signals are multiplied with different phases. The output was divided with the normalised voltage $V_b$ to cor-
rect for the non-ideal limiter. This however had a minor effect because $V_b$ only varied 4%.

Figure 6.35: Test bench output from sweeping the phase between the generators.

Sweeping the control voltage $V_c$ gave the result shown in Figure 6.36. The signal is corrected for amplitude variations which amounts to 17%. Further the signal is scaled to make it possible to compare it directly with the phase sweep in Figure 6.35. It is observed that the output voltage varies from -0.12 volt to 0.06 volt in Figure 6.36, which corresponds to a phase variation from -50 degrees to 30 degrees (see Figure 6.35). This corresponds to 80 degrees phase shift which is almost the 90 degrees predicted by the Aplac simulation. For larger values than 6 volt the phase shifter saturates.

6.4.7 Gain Stage

The gain stage is a printed circuit board with an amplifier from Mini circuits (VAM-6). The schematic is shown in Figure 6.37. The gain stage is used to clip the signal after the phase shifter to ensure a constant envelope signal (see Figure 6.14). The amplifier is specified in Table 6.4. The stage was measured as shown in Figure 6.38 and the result is depicted in Figure 6.39. The 1dB compression point was measured to -0.9dBm and the gain to 12.9dB. The gain was 2dB lower than specified and the 1dB compression point 3dB lower. Also as indicated in Figure 6.38 the bias current was set to 20mA which is 4mA more than specified. The reason for
the losses was that it was difficult to achieve a good matching with striplines of $150\Omega$ on the PCB.

The power control voltage, $V_{pc}$, is set to 3.6 volt by tuning the offset voltage in the second stage of the PCB board (see Section 6.4.5 "Error amplifier for the feed back loops" on Page 98). The

### Table 6.4: Mini Circuits Amplifier VAM-6

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gain</td>
<td>15 dB</td>
</tr>
<tr>
<td>$P_{1\text{dB}}$</td>
<td>2 dBm</td>
</tr>
<tr>
<td>NF</td>
<td>3 dB</td>
</tr>
<tr>
<td>IP3</td>
<td>14 dBm</td>
</tr>
<tr>
<td>Ibias</td>
<td>16 mA</td>
</tr>
</tbody>
</table>

#### 6.4.8 Closing the power loop

To see the improvement of linearity and to expose any stability problems of the linearisation system, the power loop was constructed and measured (Figure 6.40). The PA is fed by a generator with a constant envelope signal of 3 dBm. This is the maximum specified in the data sheet. We want the PA to work near the maximum output power level to achieve maximum efficiency. The power control voltage, $V_{pc}$, is set to 3.6 volt by tuning the offset voltage in the second stage of the PCB board (see Section 6.4.5 "Error amplifier for the feed back loops" on Page 98). The
control voltage for the desired amplitude \( A \) is set by voltage supply. The used amplitude detector is the version without a limiter (see Section 6.4.3 "Amplitude detection" on Page 95). The expected power levels and voltages are also shown in Figure 6.40.

The result of sweeping the control voltage \( A \) of the linearisation system is shown in Figure 6.41. The voltage shown was attenuated 11 dB by an attenuator and a power splitter. So the maximum voltage of 2.51 volt in 50\( \Omega \) correspond to 32 dBm at the output of the PA. This is also the maximum specified output power. The control voltage \( A \) is 160 mV for maximum output power. This is consistent with the maximum voltage coming from the amplitude detector predicted in Figure 6.40.

The characteristic of the power loop is compared with a stand alone PA (Figure 6.42). The sweep was only performed down to output voltages of 1.6 V. Hereafter is was difficult to control the output voltage manually due to the strong non-linearity from the control voltage, \( V_{pc} \), to the output power, \( P_{out} \). The two measurements are shown together in Figure 6.43. The control voltages are normalised because they cover different intervals. The curve of the linearised PA is much closer to a straight line than the stand alone PA. Further it is noticed that it was possible to control the output power to a much smaller value manually for the linearised system.

6.4.9 Conclusion on experiments with discrete RF components

The previous sections reported design and measurement of all the functional block in a polar loop type linearisation system. Several amplitude detectors were evaluated. Two printed circuit
boards, the phase shifter and the clipping amplifier, were designed, simulated and layouted. The 890 MHz phase-shifter and clipping amplifier were measured, and showed the intended functionality. Two circuit boards for subtraction of the amplitude and the phase of PA respectively were designed and tested for the proper functionality. At last a power loop was built using the designed components. It was shown that linearity was improved and that the output amplitude could be controlled over a large interval.

6.5 PA design

This section describes a CMOS power amplifier design. It is a summary of the experience collected working with power amplifier design. The main results have been published in [43][35][44] [45][46][47][48] while in the following presentation there is room for details that...
were left out in the publications. The motivation for employing CMOS technology is the low production costs and the high integration level. CMOS allows integration of control logic for all kinds of purposes with the PA’s. It is possible to switch in and out gain stages and matching circuits for different frequencies or power levels. It is possible to adjust bias for different power levels. Finally, it is also possible to employ linearisation or parts of the linearisation circuits such as AM-detectors, error control amplifiers for PA biasing.

When this work started state of the art for CMOS designs was as follows. Few class B/AB power amplifiers were published (see Table 6.5) and the achievable output power was low; approximately 19.3 dBm. There had been some publications on class E and class C power amplifiers with high efficiency and an output power of approximately 30dBm. These amplifiers are more non-linear and difficult to employ for systems with amplitude modulated signals.

In the following the design of a class B CMOS power amplifier is described. The reason for choosing class B operation is due to oxide break down considerations and because class B operation is linear compared to class C and E. The power handling capability [56] for class B PA is better than for class C and class E, which means that the output transistor is less stressed concerning maximum current and maximum voltage given a required output power. The purpose
of the design described in the following is to design a CMOS class B PA that achieves a higher output power than previously published.

### Table 6.5: Published CMOS power amplifiers

<table>
<thead>
<tr>
<th>Author</th>
<th>Power [dBm]</th>
<th>PAE[%]</th>
<th>Operation Class</th>
<th>Frequency [MHz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsai [7]</td>
<td>30.0</td>
<td>41.0</td>
<td>E</td>
<td>1980</td>
</tr>
<tr>
<td>Su [57]</td>
<td>30.0</td>
<td>42.0</td>
<td>C</td>
<td>830</td>
</tr>
<tr>
<td>Rofugaran[58]</td>
<td>13.0</td>
<td>25.0</td>
<td>C</td>
<td>1000</td>
</tr>
<tr>
<td>Yoo[59]</td>
<td>13.1</td>
<td>30.0</td>
<td>AB</td>
<td>433</td>
</tr>
<tr>
<td>Ballweber[60]</td>
<td>19.3</td>
<td>30.0</td>
<td>AB</td>
<td>900</td>
</tr>
</tbody>
</table>

### 6.5.1 The effects of low oxide breakdown voltage

The performance of a CMOS PA is limited due to supply voltage restrictions. Below the restrictions and consequences are listed.

1) The first important factor is the limitation in a CMOS process due to a low gate oxide breakdown voltage (~8V for 5nm gate oxide). The breakdown voltage limits the maximum gate-drain voltage and thereby the maximum supply voltage. Depending on the class of operation (B,C,E) of the output stage, the relation between the maximum drain voltage and the supply voltage varies. For class B operation the ratio is 2, whereas for class E the ratio is 2-5 [5]. Therefore, a class E amplifier requires a lower supply voltage than a class B PA for the same breakdown voltage.

2) The second limitation is the low output impedance of the PA. To achieve a high output power, a low effective resistive load, $R_{PA}$, (see Figure 6.44) must be chosen. The result is a high sensitivity to parasitic resistance in the matching circuits and a complicated matching circuit due to a higher transformation ratio.

3) The third limitation of the output stage is the maximum output current needed to achieve the required power in the load. The current may be so high that electromigration and parasitics in the circuit cause performance degradation.

![Figure 6.44: Power amplifier and load.](image)

The PA is designed in a 6 metal layer 0.25µm CMOS process with the data shown in Table 6.6. The breakdown voltage is approximately 8 volt. The output stage is operated in class B and the
supply voltage is 3 volt the maximum drain voltage is 6 volt. This ensures a reasonable margin before destruction of the gate oxide.

<table>
<thead>
<tr>
<th>Table 6.6: EPI-CMOS technology specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metal layers</td>
</tr>
<tr>
<td>Minimum gate length</td>
</tr>
<tr>
<td>Gate oxide</td>
</tr>
<tr>
<td>Gate oxide break down</td>
</tr>
<tr>
<td>Epitaxial layer</td>
</tr>
<tr>
<td>Substrate</td>
</tr>
<tr>
<td>Junction breakdown</td>
</tr>
</tbody>
</table>

6.5.2 PA topology

The chosen topology (see Figure 6.45) for the design is a two stage amplifier with an on-chip interstage matching consisting of the inductor, $L_1$, and a capacitor, $C$. The number of stages depends upon the necessary gain and reverse isolation. Simulations showed that two stages would suffice. The bias voltage of the gates of the transistors is delivered through resistors, $R_b$. The resistivity should be chosen significantly higher than the input impedance of the transistor to make sure the RF signal-current is led into the gate.

The inductor, $L_1$, is a part of the interstage matching. It resonates out the gate capacitances of transistor $M_2$ and parasitic capacitances $C_p$ of the capacitor, $C$. To make sure that the on-chip matching circuit is insensitive to bondwire-inductance and PCB wiring, the inductor $L_1$ is AC grounded on chip with the capacitor, $C_{\text{gnd}}$. This makes the design robust but we give up the advantage of making some off-chip tuning of the resonance frequency for the interstage matching. Examples of on-chip grounding is found in [43][44] and an example of chip tuning is found in [46].

Figure 6.45: Two stage power amplifier.

The output matching of the current PA is off-chip. It consists of the inductors $L_b$, $L_2$ and the capacitor $C_2$. The purpose of the circuit is to transform the output impedance, $R_{PA}$, of the transistor $M_2$ to the standard 50Ω used for commercial RF equipment. The higher the transformation ratio $50Ω/R_{PA}$ the more difficult the matching circuit gets in the sense that the sensitivity to
component deviation and parasitics increases. To alleviate the sensitivity more matching stages can be used in series. This also increases the bandwidth of the matching circuit. Some simple matching stages are shown in Figure 6.46. The T-section and the Pi-section are more broadband but at the cost of one extra component.

![Matching stages diagram](image)

**Figure 6.46: Matching stages (a) L-match (b) T-match (c) Pi-match.**

Not only discrete components are usable in the matching circuits but also transmission lines and stubs realized as microstrips can be utilized on a printed circuit board. One advantage of using microstrips on a PCB is that low losses can be achieved. This improves the power added efficiency of the power amplifier. An example of a design with CMOS PA and PCB-striplines where high efficiency was achieved can be found in [48].

A good explanation of passive microwave circuits can be found in [61] and a complete set of T and Pi matching circuits with equations for component values can be found in [62].

### 6.5.3 Designing the output stage

The most common way to design the output stage is to draw a load line in a drain-current versus drain-voltage plot similar to the one in Figure 3.4 on Page 35. The output transistor is then optimized along with a matching network until the right size of the transistor is found. The author instead suggest to simulate with a high-Q LC tank (see Figure 6.47) to find the right load resistance, $R_{PA}$. Subsequently the large signal drain-source capacitance necessary to design the output matching network can be found using the same circuit simulating with two different Q values. One advantage of using a high Q tank is that all harmonics of the drain current are attenuated. Secondly, the drain source capacitance can be neglected because the capacitance in the LC tank will be much larger. So it is possible to deal with load resistance separately and afterwards find the equivalent large signal drain-source capacitance.

![High Q Tank](image)

**Figure 6.47: Output transistor loaded with a high-Q LC-tank.**
The loss $R_{PA}$ in the tank is swept and the power and the PAE is plotted. For each new load, $R_{PA}$
the value of the inductor $L$ and the capacitor $C$ is recalculated so a constant quality factor is
maintained.

$$C = \frac{Q}{2\pi f_0 \cdot R_{PA}} \quad L = \frac{R_{PA}}{2\pi f_0 \cdot Q} \quad (6.22)$$

The plot of such a sweep is shown in Figure 6.48. We see that the output power and the PAE
have different maximum. The maximum for the PAE is found for higher load resistances where
a clipping takes place due to the knee voltage of the transistor. In 3-4 iterations it is possible to
find the right transistor size and load resistance with the optimal power and PAE. Since a non-
linear simulator is used for the calculation the non-linearity of the transistor is taken into account
and a realistic value of the PAE is available, especially if we want to achieve efficiency by driv-
ing the transistor into compression.

The size of the output transistor must be chosen large enough to minimize the knee voltage of
the transistor and allow a large voltage swing. The effect of the knee voltage can be seen in
Figure 6.49. At low drive levels at the gate the drain current assumes the form of the top part of
a sinusoid (left curve). At higher drive levels the drain voltage drops and the transistor enters
the triode region. Then the output impedance of the transistor drops and the transistor is not ca-
pable of draining more current. This is the reason for the bifurcated top shape of the drain cur-
cent in Figure 6.49 (figure to the right).

To monitor the state of the operation it can be useful to plot the drain current components as a
function of the gate drive level (Figure 6.50). It is seen that the fundamental component of the
drain current starts to compress at 1 volt gate drive level. This is also where the 2nd harmonic
has its maximum. From the right graph Figure 3.10 on Page 41 we see that for proper class B
operation the 2nd harmonic should be approximately half the fundamental component and the
3rd harmonic should be zero. This seems to be the case at 1.2 volt. So if we were only looking at the second and third harmonic we would identify class B operation when it has already entered light compression. But if we take the fundamental component into consideration we can avoid this mistake. So with some care it is possible to identify the class of operation based on the harmonics of the drain current. The reason that the harmonics do not behave exactly as predicted by the ideal calculations shown in Figure 3.10 on Page 41 is that the output transistor is non-linear itself and introduces harmonics.

The harmonic distortion has also been plotted. It is a convenient measure to track all harmonics at once. We see that the harmonics increases constantly after compression starts. This is due to the fact that the bifurcated drain current gets closer to zero in the middle (see right graph in Figure 6.49) which leads to higher order spectral components.

The output transistor in Figure 6.47 was simulated. The output power and the PAE was plotted similar to the graphs shown in Figure 6.48. At the load resistance of 2.5Ω a maximum output power of 1.1W and a drain efficiency of 69% was achieved. The gate drive level was 1 volt peak and a bias at the gate 0.4 volt. The supply voltage was 3 volt and the tank quality was larger than 10.

---

**Figure 6.50:** Drain currents as function of gate drive level.

**Figure 6.51:** The resonance of the circuit depends on the actual Q of the LC due to the drain-source capacitance ($f_c$=high Q resonance, $f_c$=low Q resonance).
Now that the load resistance is found we can determine the drain source capacitance. This can also be carried out with the circuit in Figure 6.47. If we lower the $Q$ value of the LC tank and recalculate $C$ and $L$ according to equation (6.22) (using the previously found value of $R_{PA}$) the drain source capacitance will change the resonance frequency of the circuit Figure 6.47 since the capacitor in the LC tank is no longer much larger than the drain source capacitance. This will change the maximum of the frequency characteristic (see Figure 6.51). Based on the new and the old maximum the effective large-signal drain-source capacitance can be calculated using the following equations.

$$
C_{ds} = \frac{C_1}{1 - \frac{R_{PA}}{2\pi f_0 \cdot X_{Cds}}} \left[ \left( \frac{f_{o1}}{f_{o2}} \right)^2 - 1 \right]
$$

where $C_1 = \frac{Q_1}{2\pi f_0 \cdot R_{PA}}$ and $r = \frac{Q_1/f_{o1}}{Q_2/f_{o2}}$ (6.23)

and

$f_0$ is the carrier frequency and intrinsic resonance frequency of the LC tanks used for both simulations.

$f_{o1}$ is the circuit resonance frequency obtained from a simulation using an LC tank with the quality factor $Q_1$.

$f_{o2}$ is the circuit resonance frequency obtained from a simulation using an LC tank with the quality factor $Q_2$.

The equation was derived by considering the circuit using two different Q-values in the LC tank. The above mentioned method was used and a large signal equivalent drain source capacitor of 9pF was found. Now a large signal model for the output transistor is available in the sense that an optimum resistive load, $R_{PA}$, and an effective drain-source capacitance, $C_{ds}$, are known. If the output matching network establishes the effective resistance and cancels the calculated effective drain-source capacitance, the required RF operation of the output transistor is ensured. The design of a matching circuit for the 50Ω load is now a simple matching exercise as shown in Figure 6.52. The transistor is modelled by a current generator, an optimal load resistor, $R_{PA}$, and the effective $C_{ds}$. The job is to match this transistor to 50Ω.

The effective $C_{ds}$ transforms the impedance in the wrong direction to lower resistance. Assuming that $C_{ds}$ is a reactance in an L-match [56] the impedance is transformed to

$$
R_1 = \frac{R_{PA}}{1 + \left( \frac{R_{PA}}{X_{Cds}} \right)^2} = 1.37 \Omega
$$

(6.24)

where $R_{PA} = 2.5 \Omega$.

To remove the reactive part of the transformed impedance an series inductor $L_b$ should be added. The inductor $L_b$ is considered as a part of an L-match and can be calculated as follows.

$$
X_{L1} = R_{PA} \cdot \sqrt{\frac{m-1}{m}} = 1.25 \Omega \quad m = 1 + \left( \frac{R_{PA}}{X_{Cds}} \right)^2 = 1.83
$$

(6.25)

$$
L_b = \frac{X_{L1}}{\omega} = 99 \cdot pH \quad \omega = 2\pi \cdot 2GHz
$$

(6.26)
Now an L-match from the impedance $R_1$ to 50Ω is calculated. The following component values are obtained.

$$L_2 = 649pH \quad C_2 = 9.50pF$$  \hspace{1cm} (6.27)

A part of the inductance $L_b$ is made up by the bondwire-inductance. If $L_b$ exceeds 649 pF the resulting impedance will be higher than 50Ω (see Figure 6.53) and another matching stage will be required to transform the impedance down. If possible the bondwire inductance should be kept below 649pH. The resulting transformation in impedance due to $L_b$ is plotted in Figure 6.53. The point where an extra matching stage will be necessary is marked.

![Figure 6.52: Matching to the large signal model of the output transistor.](image)

In principle the wanted transistor load, $R_{PA}$, the effective drain source capacitance, $C_{ds}$, and the bondwire inductance, $L_b$, determine if another matching stage is necessary. The resistance $R_2$ in Figure 6.52 seen from the 50Ω load can be calculated by equation (6.24) repeated below

$$R_1 = \frac{R_{PA}}{1 + \left(\frac{R_{PA}}{X_{Cds}}\right)^2}$$  \hspace{1cm} (6.28)

$$R_2 = R_1 \cdot \left[1 + \left(\frac{X_{L2}}{R_1}\right)^2\right]$$  \hspace{1cm} (6.29)

![Figure 6.53: Resulting transformation resistance as function of the inductance in the 2nd L-match.](image)
If \( R_2 \) is greater than 50\( \Omega \) another matching stage is necessary.

### 6.5.4 Gain stage and interstage matching

In this design a standard digital 0.25\( \mu \)m EPI-CMOS process is used. This process provides sufficient gain to achieve the desired gain of 20 dB with two stages (Figure 6.54).

In order to control the DC bias voltage of the output transistor independently of the drain voltage of the input transistor, a large capacitor, \( C \), of 20pF is inserted. The parasitic capacitance to the substrate will be significant, because the actual capacitor requires a large area as no thin oxide for the MIM capacitors is available. The on-chip inductor, \( L_1 \), resonates out the parasitic capacitance consisting of the gate capacitance of the output transistor and the parasitics \( C_p \) of the on-chip DC block capacitor. A simple but realistic inductor model derived from the current sheet model [63] was used in the design phase (see Section 6.5.6.1 "Inductor design for interstage matching" on Page 116), which made it possible to find the inductor within a few iterations. Below the inductor, a grounded poly shield [64] (see chip photo Figure 6.63 on Page 121) was placed to prevent capacitive coupling to the substrate. The shield is patterned to prevent induced currents in flowing which would lead to a reduction in the quality factor and the inductance. Losses will still be introduced in the low resistivity substrate (10\( \Omega \)-cm) due to the magnetic field, which together with the resistive wire-loss limits the Q-values to a maximum of 5 or 6. The final inductor was verified by the inductor simulator program, ASITIC, and showed agreement within 5% of the simple model.

The input transistor operates in class A and is sized to cancel the losses of the inductor and to provide a voltage swing of 1V on the gate of the output transistor (see "Designing the output stage" on Page 109). The bias voltage is provided through a 400\( \Omega \) on-chip resistor to avoid loading the AC signal on the gate.

### 6.5.5 Importance of ground bondwire inductance

The bondwire-inductance that connects chip-ground to PCB-ground has a big influence on the achievable output power. Below in Figure 6.55 a schematic of the PA and the bondwire-inductor is shown. Simulations in Figure 6.56 show that the output power decreases significantly with the bondwire inductance. To keep the power loss below 5 percent the inductance has to be kept below 0.1nH which is difficult because one bondwire alone amounts to around 1nH. So several bondwires must be connected in parallel to achieve a low bondwire inductance.
6.5.6 Layout

The layout (Figure 6.57) of the PA was in a 0.25 µm, 6 metal layer digital CMOS process. Two large capacitors are used. The DC block capacitor in the middle of the layout was realized as lateral flux capacitors in metal 3 and metal 4, whereas metal 5 and 6 were laid out as plates on top. The reason why lateral flux capacitors were used for metal 3 and 4 is that the spacing be-
between adjacent metal is smaller (0.4 µm) than the spacing between two different metal layers (0.9 µm). Utilizing metal layers 3, 4, 5 and 6, a capacitance density of approximately 0.19 fF/µm² is achieved. Metal layer 1 and 2 are not used, to limit the coupling to the substrate. The DC block capacitor is approximately 20 pF. The grounding capacitor, \( C_{\text{gnd}} \), in Figure 6.54 is realized using all metal layers. The density of this capacitor is approximately 0.39 fF/µm². The total capacitance is 46 pF. The capacitor value is estimated based on the area of adjacent metal layers and the plate capacitor formula which yields a minimum guaranteed capacitor value as fringing fields are not incorporated. The output transistor is folded to minimize the drain and source area. The maximum width of each folded transistor is 12 µm to ensure a small gate resistance and thereby a fast transistor. Also after every 5th transistor, a column of substrate contacts is inserted to ensure a low resistance to the substrate. The ground bonding inductance should be minimized using many ground pads and bonding wires. In this design 19 ground pads were used. The on-chip inductor was realized in metal layer 6 with a poly shield to reduce capacitive coupling to the substrate (see Figure 6.63 on Page 121).

### 6.5.6.1 Inductor design for interstage matching

On-chip inductors have a limited quality factor \[65][64][66][67]. This is mainly due to losses in the substrate and in the wire itself. In a normal epitaxial CMOS process, it is difficult to achieve more than seven\[66], but if a bulk CMOS process is used, a quality factor better 12 can be achieved \[66]. The difference lies in the resistivity of the substrate. An EPI-CMOS process has a low resistivity substrate of approximately 10 mΩ·cm whereas bulk processes have 1000 times higher resistivity (15-20 Ω·cm) \[66]. Therefore less power is consumed in the substrate as (eddy-) currents are induced by the magnetic fields\[68]\[67]\[64]. It seems that bulk processes gain ground again probably because the latch-up problem is controllable today. This is of course a big advantage for RF CMOS design as the substrate losses are reduced. Also, as the number of metal layers increases, the substrate losses are reduced because the distance to the substrate is increased. The tendency towards adding more metal layers is driven by the huge industry of dig-

---

**Calculation of a square spiral inductor**

Here are the design parameters set

- \( f_0 = 2 \times 10^9 \)
- Number of turns: \( N = 1.5 \)
- Inner diameter of inductor: \( d_{\text{in}} = 130 \times 10^{-6} \)
- Wire width: \( W = 30 \times 10^{-6} \)
- Wire spacing: \( sp = 3 \times 10^{-6} \)

**Results**

- \( f_0 = 2 \text{GHz} \)
- \( Q = 5.375 \)
- \( d_{\text{in}} = 130 \mu\text{m} \)
- \( d_{\text{avg}} = 176.5 \mu\text{m} \)
- \( d_{\text{out}} = 223 \mu\text{m} \)

**Inductor model parameters**

- \( L = 0.666 \text{nH} \)
- \( R_s = 1.557 \)
- \( C_{\text{OX}} = 60.373 \text{fF} \)
- \( f_{\text{SR}} = 25.1 \text{GHz} \)
- \( f_s = 10.04 \text{GHz} \)

**Equivalent parallel resonator parameters**

- \( L_p = 0.689 \text{nH} \)
- \( R_p = 46.536 \)
- \( C_p = 29.177 \text{fF} \)
- \( f_{\text{res}} = 35.496 \text{GHz} \)

---

**Figure 6.58:** MathCad interactive inductor optimization.
ital chip design. The bottleneck of the area efficiency in digital design is the inter-connectivity which is improved as the number of interconnect layers is increased.

The wire resistance is as mentioned a critical factor too. It is helped by the fact that the upper metal layers in a CMOS process are thicker and sometimes realized by materials with an extra low resistivity. Some processes even use copper in the top layer [69].

The inductor used in the current design was optimized using a math program. The program was used only to calculate the equations that will be explained in the following. No optimization or solving routines were used. It is possible to change the dimensions and immediately see the effect. The right dimensions of the inductor can be found after a few iterations. A screen dump of the window used for hand optimization is shown in Figure 6.58. The left column are the input parameters such as frequency, number of turns, wire width and inner diameter. The three columns to the right are results of the calculation. The complete math program is found in Appendix C. Now the equations and models used to model the on-chip inductor are explained. The complete inductor model [65][70] is shown in Figure 6.59. \( L \) is the inductance. \( R_s \) is the series resistance of the wire including skin effect. \( C_s \) is the capacitance that models the under path of the inductor and the coupling between the windings. \( C_{ox} \) models the coupling to the substrate. \( R_{sub} \) and \( C_{sub} \) model the substrate.

\[ L = \frac{\mu_0 N^2 d_{avg}}{2} \left[ \ln \left( \frac{c_2}{\rho} \right) + c_3 \cdot \rho + c_3^2 \cdot \rho \right] \]  

(6.30)

The parameters \( c_1, c_2, c_3 \) can have a physical meaning for rectangular and circular inductors according to the author in [63].

The inductor series resistance is based on the length, the width and the skin effect.

---

1. For more details on the current sheet model please refer to Mr. Mohan’s Ph.D. thesis.
\[ R_s = \frac{\text{len}}{\sigma \cdot W \cdot \delta \cdot \left(1 - e^{-\frac{t_{metal}}{\delta}}\right)} \]  \hspace{1cm} (6.31)

where

- \(\sigma\) is the resistivity
- \(\delta\) is the skin depth
- \(t_{metal}\) is the metal height
- \(\text{len}\) is the length of the wire
- \(W\) is the width

The total capacitance between the inductor and the substrate, \(2C_{OX}\), is calculated as a plate capacitor based on the area of the inductor. Please refer to Appendix C for more details.

Based on the PI model in Figure 6.59 an equivalent parallel resonator is calculated. This is shown in Figure 6.60.

This is convenient because the inductor is used to resonate other parasitic capacitances out in the power amplifier. The impedance at the node at resonance will be the parallel resistance, \(R_p\), calculated in Figure 6.60. The parallel resonance can also be used to predict the self resonance frequency of the inductor which is the frequency where the reactance ceases to be positive.

The inductor used in the PA design was a 0.64nH inductor with a Q value of 5. The designed inductor was verified with ASITIC and showed good agreement as mentioned earlier. The final adjustment of the inductor was carried out in ASITIC because the exact inductor value including the connecting wires should be calculated.

6.5.6.2 Capacitor design

The used capacitors are lateral flux capacitors. In this design the exact capacitance values are not needed as long as a minimum value is guaranteed. Therefore the used formulas do not account for fringing fields. They only account for the area of adjacent metals. This is indicated in Figure 6.61. The length of adjacent metals, \(L_{swall}\), is calculated based on the depicted parameters. These parameters are the width, the length and the number of fingers in one capacitor plate and the spacing between plates.

Although the lateral flux between metals is used in the same layer, the coupling between different metal layers should be utilized too. The area of overlaying metals should be calculated so the capacitance is accounted for. The used equations are as follows.
The length of adjacent metals in the same layer.

\[ L_{\text{swall}} = (L - sp) \cdot (2 \cdot N - 1) + 2 \cdot N \cdot W \]  

(6.32)

The associated capacitance is also a parameter of the height of the metal, \( t_{\text{metal}} \) and the distance between metal, \( sp \).

\[ C_{L\text{flux}} = C_{\text{swall}} \cdot L_{\text{swall}} \]  

(6.33)

where the side-wall capacitance density is

\[ C_{\text{swall}} \equiv \frac{(\varepsilon \cdot t_{\text{metal}} \cdot L)}{sp} = \frac{\varepsilon \cdot t_{\text{metal}}}{sp} \]  

(6.34)

The dimensions and area occupied on the chip.

\[ L_1 = 2 \cdot N \cdot W + (2N - 1) \cdot sp \]  

(6.35)

\[ L_2 = 2 \cdot W + L + sp \]  

(6.36)

\[ A_{\text{tot}} = L_1 \cdot L_2 \]  

(6.37)

Based on the previous figures a area density capacitance can be calculated.

\[ \frac{C_{L\text{flux}}}{A_{\text{tot}}} \equiv \frac{(\varepsilon \cdot t_{\text{metal}}/sp) \cdot 2 \cdot N \cdot L}{2 \cdot N \cdot L \cdot (W + sp)} = \frac{\varepsilon \cdot t_{\text{metal}}}{sp \cdot (W + sp)} \]  

(6.38)
If more metal layers are used the contributions of the separate layers as well as the contributions from coupling between different metal layers have to be incorporated. To be able to know when it pays off to use lateral flux capacitors a boost factor is defined. This is simply the ratio between the plate and the lateral flux capacitor. The boost factor is a function of the metal height, the distance between metal layers, the spacing and the minimum metal width. For the used technology the number is greater than one which indicates that lateral flux capacitors are more area efficient. The properties of metal layers are listed in Table 6.7.

\[
\frac{C_{Lflux}}{C_{plate}} = \frac{d_{metal} \cdot t_{metal}}{sp \cdot (W + sp)}
\]  

(6.39)

where \(d_{metal}\) is the vertical distance between the metals and \(t_{metal}\) is the metal height.

### Table 6.7: Properties and design rules of the metal layers.

<table>
<thead>
<tr>
<th>Metal #</th>
<th>Width [μm]</th>
<th>Spacing [μm]</th>
<th>Height of metal [μm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1</td>
<td>1.8</td>
<td>1.03</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1.03</td>
</tr>
<tr>
<td>4</td>
<td>0.4</td>
<td>0.4</td>
<td>0.62</td>
</tr>
<tr>
<td>3</td>
<td>0.4</td>
<td>0.4</td>
<td>0.62</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
<td>0.4</td>
<td>0.62</td>
</tr>
<tr>
<td>1</td>
<td>0.4</td>
<td>0.4</td>
<td>0.62</td>
</tr>
</tbody>
</table>

#### 6.5.7 Simulations and measurement

A printed circuit board (PCB) was designed and the bare PA-chip was glued directly on the print and bonded (chip on board) as shown in Figure 6.62. Thereby the shortest possible bond-wires were achieved. The maximum output power was measured to 29.2 dBm (see Table 6.8) at the frequency of 1.95 GHz, which to the author’s knowledge is the best performance reported for a class B power amplifier in standard digital CMOS process. In [59] and [60] an output of 13dBm and 19.3dBm respectively were achieved for a class AB PA. In [46] an output power of 31.2dBm was achieved with a CMOS process with high resistive substrate (10-20 Ω-cm) and analog options. High resistive substrate is an advantage for RF-circuits because the inductors achieve 50-100% higher Q-values and the PA-gain is higher dissipating the same amount of power. Further the process benefits from thin oxide MIM capacitors which leads to lower parasitic capacitance of the on-chip capacitors. The results achieved for the CMOS PAs [7],[57],[58] are 30dBm but these operate in class C and E which unlike class B are non-linear modes of operation. The power added efficiency of our PA was 27.4% which is approximately the same as for the two other class B CMOS amplifiers (23% in [59] and 30% in [60]).

The small signal power gain was 20.7 dB. A measured power sweep depicting the gain, PAE and output power is shown in Figure 6.62. The used frequency is the centre of the up-link frequency band of UMTS which was the design target.

The initially simulated output power was 1W and the maximum measured output power was 0.84W. The difference is due to PCB parasitics such as wire inductance and wire capacitance. Furthermore the discrete components in the matching network causes losses. The PCB wires
were modelled with strip lines and the components in the matching network with finite Q models which resulted in an output power within 5% of the measured.

6.5.8 Conclusion

A 1.9GHz class B power amplifier was designed in a digital 0.25\(\mu\)m CMOS process. The designed and fabricated PA delivered an output power of 29.2 dBm (0.84W) which to the author’s knowledge is the best result for a class B power amplifier in a standard digital CMOS process with low resistivity substrate. Agreement within 5% between measurement and simulation was achieved by carefully modelling wires and discrete components of the PCB. A design method based on deriving large signal parameters of the output transistor using an LC tank was present-
ed. The design method and the models of the passives, on-chip inductor and lateral flux capacitors were described.

Table 6.8: Measured PA characteristic.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>0.25µ CMOS</td>
</tr>
<tr>
<td>Supply voltage [V]</td>
<td>3</td>
</tr>
<tr>
<td>Frequency [GHz]</td>
<td>1.95</td>
</tr>
<tr>
<td>Max Power [dBm]</td>
<td>29.2</td>
</tr>
<tr>
<td>PAE [%]</td>
<td>27.4</td>
</tr>
<tr>
<td>small signal gain [dB]</td>
<td>20.7</td>
</tr>
<tr>
<td>large signal gain [dB]</td>
<td>14.2</td>
</tr>
</tbody>
</table>
6.6 AM-detector design

In the following chapters, the design of three amplitude detectors are presented and evaluated. The purpose is to investigate the sensitivity of different CMOS AM-detectors as only little information on this topic is found in the literature. The linearity will be investigated. Based on the measurements of the fabricated chips, the linearity of a linearised transmitter will be predicted given a certain modulation. The first circuit is a passive AM-detector with a NMOS transistor coupled as an diode. The application is discussed and simulations and measurement of a fabricated AM-detector chip is shown. The second AM-detector is based on the non-linearity between the gate voltage and the drain current. Two detectors are incorporated in the design and the detected currents of the detectors are subtracted on-chip. So this circuit delivers an amplitude error signal ready for a linearisation circuit. Again simulations and measurement of a fabricated chip is shown.

The third detector is a linear detector. It is based on a limiting amplifier and mixer. A chip has been sent to fabrication but is not available before completion of this thesis. However, the design and simulations are presented.

6.6.1 Passive AM-detector

The two circuits considered for the passive AM-detectors are based on a diode coupled NMOS transistor as shown in Figure 6.64. The left circuit has a bias current input at the drain to improve the sensitivity at low input levels. The bias current is fed by a big inductance. A capacitor in the signal path isolates the drain DC level from the external DC level. The RF signals is fed through $R_2$. The resistor $R_2$ insures a minimum load resistance to the circuit that the AM-detector connects to. In a typical application the inductor $L_1$ will be omitted and the bias is fed directly through the resistor $R_2$. This saves a lot of chip-area. The other detector considered has a separate input for the gate bias voltage. In this way a channel is created so a higher sensitivity for small input signals is achieved.

The circuits were simulated with $R_1$ equal to 1kΩ and $R_2$ equal 10kΩ. The output voltages for the circuits are shown in Figure 6.65. The bias voltages were 0.4V ($V_{bias}$) and 1V ($V_{bias,2}$). The gate biased circuit has the highest sensitivity. This can be explained in two possible ways. One explanation is that the bias condition is close to saturation, so when the RF signal is positive, the transistor will be in saturation and drain current with a high output impedance. The other
explanation is that the drain capacitance is lower due to a higher drain potential which lead to a lower drain-bulk capacitance, $C_d$. So the RF-signal is less attenuated due to the lowpass filter made up by the resistor $R_1$ and drain-bulk capacitance, $C_d$.

For the fabrication the drain biased circuit was chosen because of the small number of pins. At the time of tape out the number of pads available for the test chip were limited due to other test circuits. Therefore a low pin count were preferred. Still it is possible to measure the detector characteristic and sensitivity. Since there exists few publications on AM-detection with CMOS it is important to get some practical results. Further the precision of the MOS9 models used to model the transistors is not known for this kind of operation condition. It is expected that higher order non-linearities will influence the behaviour of the detector because it operates close to the off state. Since this is not the normal utilization of the transistor modelled, we do not know how well modelled these effects are. The normal utilization of the transistor model is for instance for LNA simulation where the signal is small and higher order effects do not have the same impact on the linearity (see Chapter 2.3). For instance the rule that the 3rd order intercept point of the harmonic distortion and the 3rd order intermodulation intercept point are proportional (see equation (2.36)) does no longer apply due to higher order non-linearities.

Before the tape-out a number of simulations were carried out to understand the detector. Some of them are discussed below.

Figure 6.65: NMOS AM-detector. a) drain biased b) gate biased.
6.6.2 Design and simulation of the fabricated AM-detector

The designed test circuit shown in Figure 6.66 can either be integrated with the PA or used separately the way it was suggested in Figure 6.8 on Page 86. The resistor $R_2$ limits the power consumed by the detector and is necessary to make sure that the main part of the power from the PA is consumed by the PA load. An on-chip LC resonator should be used to filter out the harmonics of the drain current as shown in the application Figure 6.67. This makes the amplitude detector independent of the harmonics generated by the PA. It has, however, the drawback that the signal is attenuated if the inductor quality of the on-chip inductor is to small. The equivalent parallel resistance is a function of the Q-value of the inductor, the inductance value and the frequency as given in equation (6.10) repeated below.

$$R_p = \omega L \left( Q + \frac{1}{Q} \right) \equiv \omega L Q$$  \hspace{1cm} (6.40)

For a realistic on-chip inductance of 1.2nH with a Q-value of 5 at 2GHz the equivalent parallel resistance $R_p$ yields 78Ω. Depending on how much we want to load the PA and thereby degrade the overall efficiency (see also equation (6.7) on Page 86) $R_3$ and $R_1$ in Figure 6.67 must be adjusted properly. The resistances will be chosen as a trade off between efficiency and sensitivity of the AM-detector. But with the $R_p$ equal to 78Ω it is almost unnecessary to use the resistance $R_3$ because it is negligible compared to the PA load resistance $R_{load}$ (2Ω-4Ω). Furthermore the PA will benefit from the resonator that will work as a harmonic trap. The capacitor $C_2$ in series

![Figure 6.66: Fabricated amplitude detector.](image1)

![Figure 6.67: PA with on-chip AM-detector.](image2)
with the drain resistor $R_2$ is omitted in the test circuit Figure 6.66. Therefore it is possible to set the drain bias current externally using a bias tee. The time constant can also be changed externally by adding a resistor or a capacitor parallel to $R_1$ and $C_1$.

A simulation of the drain biased circuit in Figure 6.64 with a bias voltage of 0V and 0.4V respectively is shown in Figure 6.68. We see that the sensitivity is improved with higher bias voltage. The DC current for the biased detector is less than $1\mu A$ because the bias voltage is below the threshold voltage. A simulation with enlarged drain-source areas is shown to see the effect of the drain-bulk and source bulk diodes. The simulation indicates that the sensitivity is lowered but the functionality is maintained. The bulk diodes acts as non-linear capacitances which also attenuates the RF signal.

The designed AM-detector was a part of a test chip. The chip also contained a PA and another AM-detector (see Chapter 6.7). A chip-photo is found on Page 135 and a layout of the detector is shown in Figure 6.69. The capacitor of 2pF is made by two lateral flux capacitors in parallel. The resistors are realized in a high resistive poly with a square resistance of approximately 190$\Omega$ and a width of 3.5$\mu m$ and 5$\mu m$ respectively. Minimum width is avoided to make the resistance less dependent on process variations. The transistor has a minimum gate length of 0.25$\mu m$ and a width of 10$\mu m$. The transistor was folded into 4 unit transistors to minimize the drain-source areas.
6.6.3 Measurements

For the measurements, a PCB was designed, and the chip was glued directly on the board and bonded. The schematic of the PCB is shown in Figure 6.70. The matching consisted of $47\,\Omega$ resistance and 5.6pF capacitance as DC block. The detector was measured using the setup shown in Figure 6.71. The equipment was controlled by a laptop with HP-Vee. The result of the amplitude sweeps were stored in a file and read into a math program [41]. The results were compensated for the losses in the coaxial cables. The results are shown in Figure 6.72. There are four sweeps. Two of the curves are identical. That is the sweeps with a bias voltage of 0 volt and 0.6 volt respectively. The third sweep is with a bias voltage of -0.2 volt. The fourth sweep is with the bias connection open. It is clearly seen that the detectors with the bias pin connected to the supply have a higher output voltage. But it was expected that an increased bias voltage would improve the sensitivity as seen in the simulations. This can not be observed. Only if the bias voltage is negative a lower sensitivity is observed. If the bias connection is open, so no DC current is able to flow into the circuit, a significantly reduction in the sensitivity is observed. This is expected because the DC current has to flow through the drain-bulk diode and is also seen in the simulations. It also complies with the circuits in Figure 6.2 on Page 82 which all have a return path for the DC current through the inductor. Another important observation from the measurements is that the curves tend to fall off and the slope can get negative for high ampli-
tudes. This happens for lower frequencies where the signals are less attenuated by the parasitic capacitances and a higher input signal level can be obtained. This is due to the drain-bulk and source-bulk diode. At a certain level they will start to conduct for negative RF input signals. The same phenomenon can be observed for schottky diodes [71] although it can be explained by reverse voltage break-down of the junction. For the (C)MOS transistor it is more likely to be the bulk diode because it will turn on before any channel breakdown occur. Based on the measurement it is difficult to tell the dynamic range because the precision of voltmeter limited the precision of the measurement. But based on statements published in [52][53] that FET transistor is less noisy than schottky diodes and that zero bias current leads to no 1/f noise, it is concluded that the lower bound is set by the white noise. If the detector is integrated with at PA the lower limit will be determined by substrate noise originating from the PA. The PA put out currents in the ampere range will inject large disturbances into the substrate.

6.7 Active AM-detector

The second AM-detector designed and fabricated is based on the non-linear behaviour between the gate voltage and the drain current. The detector was designed with two RF inputs and one current output as shown in Figure 6.73. The functionality of the circuit is as follows. The RF carrier at the input is led to the gate of the NMOS transistor. Due to the non-linear relationship between the drain current and the gate voltage the drain current will contain a DC component.
and harmonics of the fundamental frequency. The high frequency components including the fundamental frequency component are short-circuited through $C_5$ and $C_6$. The DC current component is mirrored through a PMOS current mirror. The current mirror contains a RC low pass filter that filters out the high frequency components. The output current, $I_{ref}$, of the left amplitude detector Figure 6.73 is mirrored and inverted through the NMOS current mirror so a current subtraction can be performed.

The harmonics generated by the detecting transistors M1 and M2 are plotted in Figure 6.74. The fundamental frequency component is the largest which means that the transistor is a good RF amplifier. The amplitudes of the even order harmonics are dominating. The 3rd and the 5th harmonics are weak in amplitude. The fundamental component and all higher harmonics are attenuated by the capacitor $C_5$ and $C_6$. The second largest component of the drain current is the DC component. This current represents the detected RF amplitude. It is also here the non-linear relation between the RF input amplitude and the output current of the detector is determined. The rest of the circuitry consist of current mirrors and RC filters which have an almost linear behaviour. The currents representing the detected amplitude is plotted in Figure 6.75. The RF amplitude at the gate of the detecting transistor is plotted too. For currents larger than 500µA the PMOS drain current is proportional to the gate voltage. The precision is within 15%. The rela-
tionship between the transistor gate amplitude and the DC drain current component is almost linear. For higher gate voltages a light compression is experienced due to the non-linear gate capacitance. The gate amplitude assumes the form similar to a square root function. The shortest signal path has been chosen for the PA input signal. In this way the pole from NMOS current mirror is avoided in the feedback loop. The reference signal goes through the NMOS current mirror without affecting the stability. The output current of the AM-detector in Figure 6.75 is multiplied by ten to see the shape clearly. This current should be zero because both RF inputs to the detector are connected to the same source as shown in Figure 6.76. Since the NMOS current mirror only mirrors the detected amplitude of RF reference signal a small attenuation is experienced for this signal. Therefore, the net output current is actually a measure of the attenuated current of the detected reference amplitude. The detector in Figure 6.73 was simulated by terminating the inputs with 50Ω as shown in Figure 6.76. The tracking bandwidth of the envelope detector is determined by the time constant made up of $C_5$ and the output impedance of the transistor M1. According to an AC simulation the output impedance of the transistor is equal to a resistor of 23kΩ ($R_t=23kΩ$) parallel with a drain capacitance of 10fF ($C_t=10fF$). If we want a 3dB cutoff frequency of 50MHz, $C_5$ must be equal to 128fF.

$$C_5 = \frac{1}{\omega_{3dB}R_f} - C_t = 128fF \quad (6.41)$$

Another time constant $\tau$ is made up of $R_f$ and $C_f$ ($\tau = R_fC_f$) in the PMOS current mirror. A pole at 500MHz is achieved with a resistance of 319Ω and a capacitance of 1pF. The poles are now placed with a distance of one decade which makes stability of a feedback loop possible.

### 6.7.1 Measurement setup

For the measurements a PCB was designed (Figure 6.77). The chip was glued directly to the PCB and bonded. The board contains resistors to match the detector to 50Ω. The detector itself has an input impedance of 1kΩ due to the on-chip resistor. Also de-coupling capacitors and the operational amplifier to convert the output current to a voltage are present on the PCB. To measure the amplitude detector the setup shown in Figure 6.78 on Page 131 was used. The generators and power supply are controlled by the computer with HP Vee. The input signals are generated by splitting the signal from a RF generator. The same generator is used to generate both input signals to ensure that the amplitude of the input signals are proportional. In this way calibrating is avoided compared to the case where an RF generator is used for each separate input. Furthermore it reflects the normal operation of the amplitude detector inside a feedback loop where the input amplitudes are approximately equal. The output current of the detector is converted to a...
voltage, $V_{out}$, by an operational amplifier. The voltage is measured by the voltmeter and stored in a file together with information about the bias voltage and the reference voltage of the operational amplifier, $V_{com}$. 

Figure 6.77: PCB test board with AM-detector.

Figure 6.78: Measurement setup.
An attenuator was used (see Figure 6.78) to compensate for different attenuation in the signal path due to mismatch. The reflection coefficient of the two RF inputs of the PCB are given in Table 6.9.

### Table 6.9: reflection coefficients of inputs of PCB2

<table>
<thead>
<tr>
<th>Input</th>
<th>Input Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1GHz</td>
<td></td>
</tr>
<tr>
<td>$</td>
<td>\rho_1</td>
</tr>
<tr>
<td>$Z = 47 - j12$</td>
<td>$Z = 46 - j9$</td>
</tr>
<tr>
<td>2GHz</td>
<td></td>
</tr>
<tr>
<td>$</td>
<td>\rho_3</td>
</tr>
<tr>
<td>3GHz</td>
<td></td>
</tr>
<tr>
<td>$</td>
<td>\rho_5</td>
</tr>
</tbody>
</table>

The gain from the RF-generator through the PA and the cables to the PA input on the PCB was 21.3dB. The gain from the RF-generator to the SMA connector at the reference input on PCB was 23.7dB.

Two different kind of sweeps were carried out. First both RF inputs were connected to the RF generator as shown Figure 6.78. Then the RF input for the reference signal was terminated with 50Ω and the RF input for the PA was connected to the RF generator. The output voltage, $V_{out}$, of the system was measured sweeping the generator amplitude and the bias voltages. Based on these measurements, it is possible to calculate the response from the RF input of the reference signal. The response from one RF input to the output does not affect the other because the output node has a fixed potential established by the operational amplifier. The output voltage can be written as follows:

$$V_{out} = V_{com} + I_{ref}R - I_{PA}R$$ (6.42)

where the currents $I_{PA}$ and $I_{ref}$ are non-linear functions of the input amplitudes.

$$I_{ref} = f_1(k_1A_{ref})$$ (6.43)

$$I_{PA} = f_2(k_2A_{PA})$$ (6.44)

The coefficients $k_1$ and $k_2$ model any attenuation from the SMA connector on the PCB to the RF input pad on the chip. The coefficients can be slightly different due to matching and trip line lengths. The difference between the coefficients can be compensated using a attenuator as shown in Figure 6.78.

From equation (6.42) we see that if the functions $f_1$ and $f_2$ are identical and the coefficients $k_1$ and $k_2$ are equal then the output will be constant as the RF amplitude is swept. From the measurement we see that this is not the case (Figure 6.79). The reason for the functions $f_1$ and $f_2$ differs is due to the fact that the NMOS current mirror does not have a infinite output impedance. So to get a close match we have to make sure that the two NMOS transistors have the same working conditions. This is achieved by choosing $V_{com}$ close to $V_{g1}$. The potential $V_{g1}$ depends on the actual current. Therefore a high bias current will minimize the relative variations of the current as the amplitude of the input signal is varying. This is also what is observed from the measurement. With a common mode voltage, $V_{com}$ of 0.625 Volt and a bias current of 700µA a good characteristic is achieved (see the blue graphs in Figure 6.79). The graph marked with squares corresponds to the measurement with signal on both inputs. It is supposed to be constant. Compared to the measurements with other bias conditions this graph was the best. It was constant for the largest input range. In the ideal case the graph should be a horizontal line cor-
responding to a constant output voltage. But for high input amplitudes the NMOS current mirror becomes imprecise due to the mismatch in the drain voltages. It should be noted that this problem can be reduced by increasing the width of the NMOS transistors. This will be shown later in Chapter 6.7.3 on Page 136 with simulation of the current mirror.

The impact of the mis-match can be calculated in the following way. Assume that the detector is inserted in a feedback loop. Then the attenuated reference signal will be compensated by reduction of the amplitude of the PA output. So the feedback will reduce the PA input amplitude until the error is zero. Two signal graphs showing the amplitude detector with and without feedback are shown in Figure 6.80. Similar to the feedback loop we can calculate how much the PA input amplitude has to be reduced to move the output current of the detector back to zero. The measured sweep with only the PA input connected to the RF generator can be used for that. This way we get the relation between the input reference amplitude and the amplitude of the PA input.
put. In Figure 6.81 the measured error is shown for equal input amplitudes (red plus). Assuming feedback, the error will be compensated by a reduction in the PA amplitude (blue diamonds). This curve is calculated using error signal (red plus). This signal drops for high amplitudes which indicates that the current representing the PA amplitude is stronger than reference amplitude (see equation (6.42)). Based on the measured response from PA input to the output a new input amplitude of the PA is calculated in such way that $V_{out}$ of the detector is unchanged. The compensated error signal (red circle) and the corresponding transfer function from the PA input to output of the detector, $A_{PA}2V_{out}$ (blue diamonds) is shown in Figure 6.81. The curves only show the corrected values. What is more important is the overall transfer characteristic of a closed loop amplitude feed back system. This characteristic can be drawn based on the amplitudes values found as the compensated values were calculated in Figure 6.81. The characteristic is shown in Figure 6.82. The slope is one which indicates a linear relation between the signal. For signals greater than 10dBm the PA amplitude compresses, which leads to distortion of the RF signal. The amount of distortion depends on the type of modulation used and the peak level (see Chapter 5.3).

### 6.7.2 Comparison between measurement and simulations

A comparison between simulation and measurement based on a amplitude sweep of one RF input while the other input is terminated with 50Ω. The blue graph in Figure 6.84 shows the measured data. The black graph is the simulated data with the same bias conditions as in the measurement. Two corrections have to be made to make the graphs overlap. The one is a scaling of input voltage to the AM-detector of a factor 1.2. This can be explained by losses in the signal path of the RF signal at the PCB. A perfect match could not be achieved. This was confirmed by the measurement of the reflection coefficient of -12dB. The other correction that must be...
made on the simulated data is adding an offset of 0.5V. This corresponds to an offset current of 194\(\mu\)A out of the detector. One explanation could be that a high bias current of 700\(\mu\)A is floating through the detecting transistors and that the current originating from the detector connected to the reference input is mirrored through the non-optimal NMOS current mirror. The current out of the mirror is reduced which leads to a net current out of the detector. But we saw in the measurement that the mirror apparently only showed weakness for high currents (several mA) which could be simulated too. This offset for low amplitudes (and low drain currents) can not be simulated. So this is not likely to be the reason. Another possibility is that the PMOS transistors are poor in the sense that they have a low output impedance. This was confirmed by an-

---

**Figure 6.82: DC detector characteristic:** \(A_{ref}\) to \(A_{pa}\).

**Figure 6.83: Chip photo of die with 2 AM-detectors and a PA.**
other designer using the same process. He only had problems with the circuits using PMOS [72]. The impedance $1/g_{M7}$ that transistor M5 looks into in Figure 6.73 on Page 128 is higher than the impedance of zero Ohm that M6 looks into. M6 looks into virtual ground of the external operational amplifier (see Figure 6.76) This can lead to an offset because of the lack of symmetry.

### 6.7.3 Study of the NMOS current mirror

The reason for missing symmetry between the RF reference input and the PA input was due to the NMOS current mirror, M7 and M8 (Figure 6.73 on Page 128). M8 leaves saturation and goes into triode operation. The symmetry can be significantly improved by increasing the width of the transistors. The circuit in Figure 6.85 was simulated with different transistor width. The amplitude of the RF input signals $S_{\text{ref}}$ and $S_{PA}$ was swept and the output voltage of the operational amplifier was plotted. Two modified detectors with NMOS cascodes shown in Figure 6.86 were also simulated to see if further improvement could be achieved. The result is shown in Figure 6.87. The red curve (cross) represents the amplitude detector with 30u transistor width and no cascode. Compared to a detector with 300µm (graph w. blue plus) in the current mirror a significant improvement is achieved. The output voltage of the detector is supposed to stay constant as the amplitudes of each RF input are equal. For low amplitudes the output voltage is equal to the voltage, $V_{\text{com}}$, of the plus terminal on the operational amplifier.

![Figure 6.85: Principle schematic used for simulation.](image-url)
This is expected due to the virtual ground principle. For higher amplitudes the current mirror ceases to mirror the current because M8 goes out of saturation. So the current coming from the M5 representing the amplitude detected from the RF reference signal appears smaller at the output of the detector. That result in a net current out of the amplitude detector which would result in a decreasing voltage at the output of the amplifier. If the transistors are increased to 300µ we see a significantly improvement (blue plusses in Figure 6.87 and Figure 6.88). Transistor M8 stays in saturation and the current is mirrored properly. It is also seen that if a cascode is used the performance can be improved. For cascodes sized to 800µm we see an overall improvement whereas a cascode with a width of 600µm performs worse than a stage without cascodes for high amplitudes.

### 6.7.4 Performance prediction of AM-detector

The different variations of AM-detectors in the previous section were supposed to stay at constant output when the same signal was applied to both input. The deviation as shown in Figure 6.88 led to a non-linear characteristic of the overall linearisation system similar to the one shown in Figure 6.82 on Page 135. Now the same principle is used to calculate the spectral
leakage and the error vector magnitude of a GSM/EDGE signal led into a polar linearisation system based on variations of the AM-detectors. First the AMAM characteristic of the envelope feedback system is calculated as described in Section 6.7.1 and Figure 6.82. In Figure 6.89 it is seen that the AM-detector with a NMOS current mirror with transistor widths of 30\(\mu\)m has a small linear input range. As soon as the width is increased to 300\(\mu\)m the linear input is improved significantly. The AMAM characteristic looks like a straight line (blue line). This is also the case for the AM-detector were NMOS cascodes were used (magenta line).

To relate these AMAM characteristics to the distortion measure used in a radio system the error vector magnitude and the spectral leakage is calculated. In Figure 6.90 the AM-detector with the narrow NMOS transistors in the current mirror has the highest EVM. The limit of 9% is exceeded at 0.21 of the maximum simulated input amplitude. The two other versions of the AM-detector respect the maximum limit for all simulated amplitudes. The spectral leakage is in agreement with what was found in Section 5.3 "Impact of non-linearities on QPSK modulated signal" on Page 73 the hardest requirement. The ACP is plotted in Figure 6.91 with a frequency offset of 200kHz and 400kHz respectively. The requirements at the frequencies are -40dBc and -60dBc respectively. It is seen that the detector with the narrow NMOS transistor in the current mirror only comply to the specifications for very small amplitudes corresponding to 0.15 of the maximum simulated input amplitude. The detector with a wide NMOS transistor in the current mirror (blue graph) complies until the amplitude exceeds 0.55 of the maximum simulated input.
amplitude. At this point the ACP at 400kHz is exceeded. Only the detector with NMOS cascodes complies with the requirements for all simulated amplitudes.

Figure 6.90: Error vector magnitude of closed loop system.

Figure 6.91: Spectral leakage.
6.8 Linear AM detector

A linear envelope detector can be designed as a limiting amplifier and a mixer as shown in Figure 6.6 and repeated below.

The RF signal is led into a limiting amplifier and a square wave is achieved. The square wave can be perceived as the sign of the RF signal. Multiplying this signal with the RF signal itself a rectified signal of the RF signal is achieved. The rectified signal is low pass filtered to obtain the envelope signal.

This is also the principle for the envelope detector in Figure 6.93. The limiter has a differential output that control a switching bridge of four NMOS transistors. The bridge changes the sign of the RF input-signal. The RF input signal is converted from a single ended signal to a differential signal using two trans-conductance stages. The lowpass filtering is carried out by three capacitors. One capacitor is coupled differential and two are coupled to filter any common mode RF signal. To make measurements simple and to accommodate single ended power amplifiers, which is the normal way to realize PA’s today, the input to the envelope detector was chosen to be single ended. On-chip, the signal is converted to differential signal. The mixer bridge requires a differential signal to be able to change sign of the RF signal. The differential signal is generated by two trans-conductance stages as shown Figure 6.94. The the inverting trans-conductance stage is a simple NMOS transistor stage. The load is a PMOS transistor with an operational amplifier to control the DC voltage of the output. The DC voltage is sensed through a resistor of 10kΩ to make sure that the RF signal is not attenuated. The gain of the operational amplifier has been set to one by resistive feedback. This was necessary to insure stability of the DC level.
control loop. Even with the low gain operational amplifiers an external capacitor at the output of the operational amplifiers of 350pF was necessary to guarantee a phase margin of 67 degrees and a gain margin of 17dB.

The non-inverting trans-conductance was realized using a PMOS current mirror to invert the current. The principle for the DC level control is similar to the inverting trans-conductance stage. The limiter is the most critical components in the design because the precision of the envelope detector is dependent of the AMPM conversion of the limiter. But also a lower limit of the envelope of the input signal has to be guaranteed to insure correct operation of the limiter. The limiter in Figure 6.95 consist of gain stages, limiting amplifiers and differential stages that converts the single ended limited RF signal to a differential signal.

The reason that the first stage is an amplifier with no amplitude limiting element is that we want to have all signal levels amplified to be large signals before they are led into a limiting stage. If we can insure that a large interval of input amplitudes experiences the same number of limiting stages we achieve a constant delay for a large interval of amplitudes. Thereby a low AMPM conversion is achieved. On the other hand if a very small signal is led into a series of limiting stages the first couple of stages will only amplify the signal and not do any clipping of the signal. The small signal will therefore experience different delay than a large signal because it has propagated though a smaller number of stages where clipping has been performed. The clipping process has an associated delay because of the non-linear capacitances in the diodes.

Two types of gain stages were considered in the design phase. One with resistive load and one with and LC tank load. First the stage with one resistor was considered because it occupies a little area. The wanted bandwidth of the gain stage could be achieved by choosing the resistive load small enough. But the problem is that the gain increases as the frequency is lowered. That means that we only got the wanted gain at high frequency (e.g. at 2GHz) but at low frequencies (e.g. 10MHz) we have a much higher gain because the effect of the parasitic capacitances have
So a small noise signal at 10MHz will be amplified with a much higher gain that a signal at 2GHz. This can lead to lead to and unwanted clipping of a noise signal instead of the wanted signal. This is also what happened in the preliminary simulations. A clipping of a low frequency signal drown out the wanted signal a 2GHz. This shows that frequency selection is necessary to insure the right frequencies are amplified. After the load was changed to an LC tank the wanted amplitude limitation was achieved.

The envelope detector was simulated in Cadence with Spectra RF [73] but as more limiting stages were added the simulator got convergence problems using steady state simulations (shooting method.). Therefore the functionality was tested with few limiting amplifier stages. Then all stages were included and some normal transient simulations were carried out with different input RF amplitudes. After settling the output voltage was used to plot an output characteristic of the detector as shown Figure 6.96. The red circles are the output values. It is seen that this detector is very linear compared to the previous. Its limitation is for small input amplitudes because the limiter ceases to operate as intended and start working as a amplifier. The linearity was calculated to 0.2% RMS based on the simulated points. A chip-photo of the designed detector is shown in Figure 6.97.

**Figure 6.96:** Output voltage of AM-detector and fitted line.

**Figure 6.97:** Chip-photo of the linear AM-detector.
6.9 Resume of chapter 6

In this chapter three amplitude detectors were presented and evaluated. The reason for the investigation was that there existed few publications on AM-detectors. Especially in CMOS technology this is true. Therefore two test circuits, one with a passive and one with an active detector, were fabricated and measured. Two types of passive AM-detectors with different bias methods were considered. The first one\(^1\) had the drain and the gate directly connected. The second one\(^2\) the drain and the gate were only AC connected so it was possible to set the gate and the drain voltage independently. From the simulations it was shown that if the drain voltage is chosen higher than gate voltage a higher sensitivity is achieved. This was caused by a operation close to saturation. In addition a smaller drain bulk capacitance was achieved. Although the drain-gate independent bias method yields higher sensitivity, the first type of detector was chosen due to it simplicity in the sense that it had less capacitors and bonding pads. The detector had a limited number of pads available due to other test circuits on the chip. The measurements yielded a characteristic similar to the simulated but with a lower sensitivity. Further two phenomena that could not be simulated were found during the measurement.

1) At high input amplitudes which was achieved at lower frequencies around 1GHz at negative slope of the characteristic was measured. This is similar to what can be experienced with Schottky diodes due to reverse conduction due to break down. But for the CMOS technology is likely to be the drain-bulk diode that start to conduct current for the negative parts of the RF signal. So it is important to be aware of the effects if the detector is used with large RF amplitude signals.

2) The effect of different bias currents could not be measured. The same sensitivity was measured for different bias currents. Still the bias was important because if the bias connection was opened a significantly sensitivity reduction was measured. This was explained by the DC return current in the detector has to flow though the drain-bulk diode.

The second detector fabricated is an active detector capable of detecting two RF signals and subtract the envelope signal-currents. The output is a current that is easily converted to a voltage with a operational amplifier. This amplifier can be integrated on the chip in a later design but was left out to minimize the risk and keep the complexity low. The non-linearity used for the envelope detection was the gate-voltage drain-current relation. The detector can be biased to increase the sensitivity. A test-bench controlled by a computer with HP-Vee was made. The fabricated chip was glued and bonded directly to PCB. Measurements under different bias conditions were carried out. The characteristic of the detector was measured. The non-linearity of the detector was higher than expected. This was due to a too small dimensioned current mirror which degraded the linearity of the detector. This non-linearity could also be simulated. Furthermore it could be shown that linearity could be significantly increased with a larger current mirror. A comparison between simulation and measurement were made. Two corrections had to be made to get a good agreement between measurements and simulations. A offset current out of the detector had to be added which is likely to be caused by a low output impedance of the PMOS transistor. The second correction was a reduction of the simulated input signal which can be explained by losses in the matching on the PCB.

To give a realistic picture of the achievable linearity of the detector, the modulation quality and the spectral leakage using a modulated signal were calculated.

---

\(^1\)The first type of passive detector is called drain-bias in the current text.
\(^2\)The second type of passive detector is called gate-bias in the current text.
Conclusion

This thesis has treated topics concerning linearisation of RF power amplifiers using feed-back with focus on polar modulation feed back. The thesis contains a chapter devoted to theoretical aspects such as architecture, design and simulations considerations. A second chapter treated important building blocks and their integration for a polar modulation feed back loop such as power amplifier, phase shifter and envelope detectors. The work carried out and the results achieved in this project are emphasized below.

• A new architecture for polar linearisation of RF power amplifiers suitable for digital transmitters was presented. The number of power consuming analog signal processing blocks were reduced and moved to the digital domain. Furthermore, an extension to the new architecture that reduces the stability constraints and maintains a constant loop-gain and system bandwidth was introduced. Simulations showed the reduction in the amplitude error was more than 6% for low amplitude levels.

The distortion sensitivity to delay difference in the signal path of the phase and amplitude was analysed, and simulations showing the relation between delay and spectral leakage were presented. Design guidelines for the envelope feed back loop was given. Based on the fact that the polar loop can be treated as two separate feed back loops simulation guidelines which lower the simulation complexity were given. The simulations based on a radio system (GSM/EDGE) that supports amplitude and phase modulation were presented. Signal quality and spectral regrowth were simulated based on a measured power amplifier. Equations that determine the settling of the Polar and Cartesian loop were found and the difference concerning the phase of the loop-gain were discussed.

• Experiments with a QPSK modulated signal were carried out to test the sensitivity to different non-linearities. Phase and amplitude non-linearities are equal severe. Furthermore it was seen that the most difficult requirement to satisfy was the spectral leakage requirements compared to in-band modulation quality.

• Three amplitude detectors were presented and evaluated. The reason for the investigation was that there existed few publications on AM-detectors. Especially in CMOS technology this is true. Two test circuits, one with a passive and one with an active detector, were fabricated and measured. Two types of passive AM-detectors with different bias methods were considered. The first one had the drain and the gate directly connected. The second one, the drain and the gate were only AC connected so it was possible to set the gate and the drain voltage independently. From the simulations it was shown that if the drain voltage is higher than the gate voltage a higher sensitivity is achieved. This was due a operation close to saturation, and due to a smaller drain bulk capacitance. The measurements yielded
a characteristic similar to the simulated but with a lower sensitivity. Furthermore two phe-
nomenons that could not be simulated were found during the measurement.

1) At high input amplitudes a negative slope of the characteristic was measured. This is
similar to what can be experienced with Schottky diodes due to reverse conduction
due to break down. But for the CMOS technology this is likely to be the drain-bulk
diode that start to conduct current for the negative parts of the RF signal. So it is
important to be aware of the effects if the detector is used with large RF amplitudes
signals.

2) No effect of changing the bias currents could not be measured. The same sensitivity
was measured for different bias currents. The same sensitivity
was important because if the
bias connection was opened a significantly sensitivity reduction was observed. This
was explained by the DC return current has to flow through the drain-bulk diode in the
detector.

• The second detector fabricated was an active detector capable of detecting two RF signals
and subtract the envelope signal-currents. The non-linearity used for the envelope detection
was the gate-voltage drain-current relation. The detector can be biased to increase the sen-
sitivity. A test-bench with a PCB and the fabricated chip on-board was constructed. Meas-
urements under different bias conditions were carried out. The symmetry of the detector
was lower than expected due to a NMOS current mirror with too narrow transistor widths.
This lack of symmetry could also be simulated and it could be shown that it could be sig-
ificantly increased with a larger current mirror. A comparison between simulation and
measurement were made. There were good agreements except for an offset current out of
the detector which could be explained by a bad performance of the PMOS transistor, and
an attenuation at the PCB input due to reflection. To give a realistic picture of the achieve-
able linearity of the detector the modulation quality and the spectral leakage were calcu-
lated based on the measurements of the detector characteristic.

• A 1.9GHz class B power amplifier was designed in digital 0.25µm CMOS process. The
designed and fabricated PA delivered an output power of 29.2 dBm (0.84W) which is the
best result for a class B power amplifier in a standard digital CMOS process with low resis-
tivity substrate. Agreement within 5% between measurement and simulation was achieved
by carefully modelling wires and discrete components of the PCB. A design method based
on deriving large signal parameters of the output transistor using an LC tank was presented.
The design method and the models of the passives, on-chip inductor and lateral flux capac-
itors were described.
Literature


[8] Changsik Yoo, Qiuting Huang, "A common-Gate Switched 0.9W Class-E Power Amplifier with 41% PAE in 0.2µm CMOS", 2000 Symposium on VLSI Circuits, pp.56-57.


[37] Patent filed in Great Britain, GB 0030684.5.

[38] S. Wong, S. Luo, L. Hadley, “A 2.7-5.5 V 0.2-1 W BiCMOS RF driver amplifier IC with closed-loop power control and biasing", *ISSCC Digest of Technical Papers*, Feb. 1998, pp.52-53.


[73] Cadence Design Systems, Inc, Analog Artist design system, Spectre RF circuit simulator, Virtuoso layout tools, Diva verification.
Index

Numerics

1-dB compression point ........................................................................................................ 22

A

ACP ........................................................................................................................................ 18, 68
Active AM-detector .................................................................................................................. 128
Adjacent Channel Power ........................................................................................................ 18
AM (Amplitude modulation) .................................................................................................. 11
AM detector design ............................................................................................................. 123
AM-detector ......................................................................................................................... 81
AM-modulated carrier .......................................................................................................... 83
amplitude detection ............................................................................................................. 85, 95
amplitude detector .............................................................................................................. 66
amplitude modulation .......................................................................................................... 11
Amplitude non-linearity ....................................................................................................... 75
analytical signal ................................................................................................................... 8
ASITIC .................................................................................................................................. 114

B

Band limited RF signal ........................................................................................................... 7
Bandpass filter ....................................................................................................................... 87
bandwidth parameter of RC-filter ....................................................................................... 15
BER ....................................................................................................................................... 18
biased peak detector ........................................................................................................... 82
Bit Error Rate ...................................................................................................................... 18
boost factor, capacitors ....................................................................................................... 120

C

CALLUM .............................................................................................................................. 47
capacitance density ............................................................................................................. 119
Capacitor design .................................................................................................................. 118
carrier signal ....................................................................................................................... 7
Cartesian Modulation Feedback .......................................................................................... 48
Chebyshev transform .......................................................................................................... 29
chip photo ............................................................................................................................. 121
class (A,B,C), operation mode ........................................................................................... 33
class B .................................................................................................................................. 36
Class C .................................................................................................................................. 37
complex envelope ............................................................................................................... 8
Complex envelope representation ...................................................................................... 8
complex envelope signal ................................................................................................... 28
complex exponential function ......................................................................................... 16
Complex Gain Predistortion ............................................................................................. 53
complex modulation function ......................................................................................... 8
INDEX

complex numbers representation ............................................................... 8
complex signal ........................................................................................................ 6
compression point (n-dB) .................................................................................... 26
conduction angle ................................................................................................. 37
constellation points .............................................................................................. 16
convolution .............................................................................................................. 9
coupling between FB loops .................................................................................. 67
current sheet model ............................................................................................... 117

d
 describing function .......................................................................................... 29
design considerations for envelope feed back ...................................................... 86
Design guide lines for a envelope feed back system ............................................. 69
Digital Predistortion ............................................................................................ 51
digital predistortion ............................................................................................... 51
Diode ..................................................................................................................... 81
diode bias-current .................................................................................................. 83
Dirac pulse .............................................................................................................. 18
direct conversion transmitter .................................................................................. 3
drain peak current .................................................................................................... 40
Dynamics of the polar loop linearisation ............................................................... 65

E
eddy-current ........................................................................................................... 116
EDGE modulation ............................................................................................... 16
EDGE modulation definition ................................................................................. 18
efficiency .................................................................................................................. 42
Envelope feed back ............................................................................................... 50
Envelope feed back system with non-linear AM-detectors .................................. 88
Envelope Feedback ............................................................................................... 50
EPI-CMOS process ............................................................................................... 116
Error Vector Magnitude ....................................................................................... 18
EVM ......................................................................................................................... 18

F
fabricated AM-detector ......................................................................................... 125
Feed Forward Linearisation ................................................................................. 51
feed forward predistortion .................................................................................... 51
Field Effect Transistor AM detectors ................................................................... 85
Filtering of complex envelope signals .................................................................. 9
Fourier expansion ................................................................................................. 28

G
Gain compensation ............................................................................................... 69
gain ideality factor ............................................................................................... 56
Gauss pulse ............................................................................................................. 18
ground bondwire inductance .................................................................................. 114
| H | Harmonic distortion ........................................................................................................... 22 |
| I | image frequency ................................................................................................................ 6 |
| I | impact of non-linearities on QPSK modulated signal .......................................................... 73 |
| I | impulse response ............................................................................................................. 9 |
| I | inductor ............................................................................................................................ 116 |
| I | Inductor design .................................................................................................................. 116 |
| I | inductor series resistance ................................................................................................. 117 |
| I | injected phase distortion ................................................................................................. 66 |
| I | inphase and quadrature conversion .................................................................................. 1 |
| I | inphase component .......................................................................................................... 7 |
| I | inphase signal .................................................................................................................. 4 |
| I | Intercept point ................................................................................................................... 26 |
| I | intermediate frequency ................................................................................................. 5, 6 |
| I | Intermodulation ............................................................................................................... 23 |
| I | intermodulation ............................................................................................................... 27 |
| I | Intermodulation intercept point ..................................................................................... 23 |
| I | intermodulation intercept point ..................................................................................... 27 |
| I | interstage matching ....................................................................................................... 114 |
| I | intersymbol interference ................................................................................................. 14 |
| I | IQ constellation plot ....................................................................................................... 15 |
| I | IQ modulation ................................................................................................................... 11 |
| I | IQ-trace diagram .............................................................................................................. 16 |
| I | ISI .................................................................................................................................. 14 |
| K | key building blocks ......................................................................................................... 81 |
| L | large-signal drain-source capacitance ............................................................................. 112 |
| L | lateral flux capacitor ...................................................................................................... 118 |
| L | Layout .............................................................................................................................. 115 |
| L | limiter ............................................................................................................................... 30, 95 |
| L | LINC ................................................................................................................................. 46 |
| L | LINC Technique ............................................................................................................... 46 |
| L | Linear AM detector ........................................................................................................... 85, 141 |
| L | Linear Amplification with nonlinear components ............................................................. 46 |
| L | linear region ..................................................................................................................... 84 |
| L | Linearisation Technique ............................................................................................... 45 |
| L | linearised Gauss pulse ................................................................................................. 17 |
| L | LO ................................................................................................................................ 4 |
| L | local oscillator ................................................................................................................ 4 |
| L | loop bandwidth ............................................................................................................... 68 |
| L | loop gain compensation ................................................................................................. 70 |
INDEX

Loop gain problem ................................................................. 67
Loop settlement ........................................................................ 64

M
Mapping Predistortion .......................................................... 52
matched AM-detectors .......................................................... 90
MathCad interactive inductor optimization .......................... 116
maximum filter gain .............................................................. 67
maximum output power ......................................................... 40
Memory-less bandpass non-linearity ................................. 28
Mini Circuits .......................................................................... 95
minimum loop bandwidth ...................................................... 68
modulation index ................................................................. 83

N
negative peak detector ............................................................ 82
NMOS current mirror, study .................................................. 136
Non-linear system ............................................................... 21	non-linearity with hysteresis ................................................. 31
Nyquist filter ........................................................................ 14

O
offset loop transmitter ............................................................ 6
on-chip AM detector .............................................................. 125
output transistor ................................................................. 116

P
PA isolation .......................................................................... 54
PA load resistance ............................................................... 112
Passive AM detector ............................................................ 123
PCB photo of power amplifier .............................................. 121
PCB photo, AM-detector ..................................................... 127
peak detector ....................................................................... 82
Performance prediction of AM-detector ............................ 137
Phase Correcting Feedback .................................................. 50
Phase correction feedback ................................................... 50
phase distortion ................................................................. 66
phase locked loop ............................................................... 6
Phase non-linearity ............................................................. 74
Phase shifter ....................................................................... 100
PLL .................................................................................. 6
Polar loop ........................................................................... 63
Polar loop linearisation using a VCO ................................. 62
Polar loop with an RF-phase shifter .................................... 61
Polar modulation feedback ................................................. 60
Polar Modulation Feedback ................................................ 48
Polar Modulation Feedback Loop ...................................... 48
Polar Predistortion ............................................................. 53
Polar predistortion ............................................................. 53
Polynomial non-linearity ................................................................. 24
Polynomial predistortion ............................................................... 54
Power Amplifier ........................................................................... 33
power amplifier ........................................................................... 4
power consumption, class C PA .................................................. 39
power control ............................................................................ 44
power efficiency ......................................................................... 42
pre-envelope ............................................................................... 8

Q
QPSK modulated random signal ................................................... 16
QPSK modulation ......................................................................... 13
quadrature component ................................................................ 7
quadrature signal ......................................................................... 4

R
raised cosine filter ........................................................................ 14, 15
RF frontend .................................................................................. 4
RF power amplifier ...................................................................... 33
Ritz-Galérkin method .................................................................. 83
root raised cosine filter ............................................................... 16

S
saturation current ......................................................................... 83
signal quality ................................................................................ 18
Simulating the polar loop ........................................................... 67
simulation model, bandpass non-linear ...................................... 29
Simulation of the envelope FB system with gain control ............. 71
single side band .......................................................................... 4
single side band modulation ......................................................... 11
small signal gain ......................................................................... 22
spectral leakage ........................................................................... 68
square law characteristic ............................................................ 84
SSB ............................................................................................... 4
substrate-resistivity .................................................................... 116
Super heterodyne transmitter ..................................................... 5
symbol ......................................................................................... 13
Symbol mapping table for EDGE ................................................ 17

T
Transmitter architecture .............................................................. 3

V
VCO .............................................................................................. 7
voltage controlled oscillator ....................................................... 7
voltage multiplier ........................................................................ 82
INDEX

W
WB-CDMA ....................................................................................................................... 10
weighted sum of sinusodials .......................................................................................... 21
Appendix A
Inphase and quadrature frequency translation

In the following the principle for inphase and quadrature frequency conversion is described. Also effects such as the gain and the phase error is described.

A block diagram for a inphase and quadrature up-conversion mixer is shown in Figure A1. Two continuous time signals $I_1$ and $Q_1$ go into a mixer here modelled by a multiplication. The signals are multiplied with a cosine signal and a sine signal with the frequency, $\omega$, much higher than the bandwidth of the signals $I_1$ and $Q_1$. The output of the mixers are subtracted and represented by the signal $S$, which is the modulated RF signal.

\[
S(t) = I_1(t) \cdot \cos(\omega t) - Q_1(t) \cdot \sin(\omega t) \quad (A1)
\]

Where $\omega$ is the carrier frequency.

By the down-conversion, the inphase signal $I_1$ and the quadrature signal $Q_1$ are extracted. Multiplying with $\cos(\omega t)$ and $\sin(\omega t)$ (see Figure A2) respectively extracts the wanted signals at low frequency again.
After multiplying the RF signal $S(t)$ with $\cos(\omega t)$ one achieves a low frequency component proportional to $I_1$ and a sum of two spectral components at the frequency $2\omega$ representing $I_1$ and $Q_1$.

$$a(t) = s(t)\cos(\omega t) = \frac{1}{2}[I_1(t) + I_1(t)\cos(2\omega t) - Q_1(t)\sin(2\omega t)]$$  \hspace{1cm} (A2)

The high frequency spectral components are eliminated by a lowpass filter and the inphase signal is obtained.

$$Q_2(t) = \frac{1}{2} \cdot I_1(t)$$  \hspace{1cm} (A3)

Similar to the inphase signal path the quadrature signal is extracted by multiplying by $\sin(\omega t)$.

$$b(t) = S(t)(-\sin(\omega t))$$
$$\quad = \frac{1}{2}[Q_1(t) - I_1(t)\sin(2\omega t) - Q_1(t)\cos(2\omega t)]$$  \hspace{1cm} (A4)

$$Q_2(t) = \frac{1}{2}Q_1(t)$$  \hspace{1cm} (A5)

This was the ideal case where no phase and gain imbalances were present. Now consider the case where we have several phase imbalances. First we have imbalance of $\alpha$ in the upconversion. The output of the upconversion will look as follows.

$$S(t) = I_1(t)\cos((\omega t) - Q_1(t)\sin(\omega t + \alpha))$$
$$\quad = [I_1(t) - Q_1(t)\sin(\alpha)]\cos(\omega t) - [Q_1(t)\cos(\alpha)]\sin(\omega t)$$  \hspace{1cm} (A6)

Note that the phase error distorts the signal in a way that does not change the bandwidth of $s(t)$. Only the phase of the quadrature signal is distorted. If the imbalance reaches 90 degrees $s(t)$ can

be considered as the difference of the inphase and the quadrature signal modulated on the same carrier.

\[ \alpha = \frac{\pi}{2} \Rightarrow s(t) = [I_1(t) - Q_1(t)] \cos(\omega_0 t) \]  
\[ (A7) \]

In this case there is no guarantee that the signals can be separated again. In the following we assume that the phase imbalance is relative small e.g. less than 5 degrees.

In the down conversion mixer we assume a phase error \( \gamma \) due to delays such as the channel. The phase imbalance inside the down conversion mixer is set to \( \beta \). After the multiplication the signal \( a(t) \), \( b(t) \) (see Figure A2) has the following terms.

\[ a(t) = \frac{1}{2} [I_1(t) \cos(\gamma) + Q_1(t) \sin(\gamma - \alpha)] \]  
\[ + \frac{1}{2} [I_1(t) \cos(2\omega_0 t + \psi) - Q_1(t) \sin(2\omega_0 t + \gamma + \alpha)] \]  
\[ (A8) \]

\[ b(t) = \frac{1}{2} [-I_1(t) \sin(\gamma + \beta) + Q_1(t) \cos(\gamma + \beta - \alpha)] \]  
\[ + \frac{1}{2} [-I_1(t) \sin(2\omega_0 t + \psi + \beta) - Q_1(t) \cos(2\omega_0 t + \gamma + \alpha + \beta)] \]  
\[ (A9) \]

Only the low frequency components are of interest. The high spectral components are eliminated by a filter. The inphase and quadrature component at the output of the downconversion mixer have the form

\[ I_2(t) = \frac{1}{2} [I_1(t) \cos(\gamma) + Q_1(t) \sin(\gamma - \alpha)] \]  
\[ (A10) \]

\[ Q_2(t) = \frac{1}{2} [-I_1(t) \sin(\gamma + \beta) + Q_1(t) \cos(\gamma + \beta - \alpha)] \]  
\[ (A11) \]

Gain error in the mixers are easily taken into account. E.g if the up-conversion mixer has a gain in the inphase path \( I_1(t) \) must be scaled with this factor in equation (A10) and (A11). If the gain error is in the inphase path of the downconversion mixer \( I_2(t) \) must be scaled with this factor. So all phase and gain errors can be dealt with in baseband. The imperfections can also be formulated matrix vector equations. Assume that \( I_1(t) \) and \( Q_1(t) \) constitute a vector. Each separate phase imbalance can den be formulated as a matrix multiplication.
\[
\begin{bmatrix}
I_2(t) \\
Q_2(t)
\end{bmatrix} = \frac{1}{2} \cdot B \cdot G \cdot A \cdot \begin{bmatrix}
I_1(t) \\
Q_1(t)
\end{bmatrix}
\]  
(A12)

where

\[
B = \begin{bmatrix}
1 & 0 \\
-\sin(\beta) & \cos(\beta)
\end{bmatrix} \quad G = \begin{bmatrix}
\cos(\gamma) & \sin(\gamma) \\
-\sin(\gamma) & \cos(\gamma)
\end{bmatrix} 
\]  
(A13)

\[
A = \begin{bmatrix}
1 & -\sin(\alpha) \\
0 & \cos(\alpha)
\end{bmatrix} 
\]  
(A14)

\[
B \cdot G \cdot A = \begin{bmatrix}
\cos(\gamma) & \sin(\gamma - \alpha) \\
-\sin(\gamma + \beta) & \cos(\gamma + \beta - \alpha)
\end{bmatrix} 
\]  
(A15)

\[B, G, A\] corresponds to affin mappings where \(G\) is a pure rotation by \(\gamma\) radians. The matrix equations (A12) can be verified by multiplying the matrices and simplifying the trigonometric relations.

This result corresponds to the equations (A10) and (A11).
Appendix B
Puls shape filter for EDGE

Pulse shaping

The modulating 8PSK symbols $\delta_j$ as represented by Dirac pulses excite a linear pulse shaping filter. This filter is a linearised GMSK pulse, i.e. the main component in a Laurent decomposition of the GMSK modulation. The impulse response is defined by:

$$c_0(t) = \prod_{\xi=0}^3 S(t+iT), \text{ for } 0 \leq t \leq 5T$$
$$= 0, \text{ else}$$

where

$$S(t) = \begin{cases} 
\sin(\pi \int_0^t g(t')dt'), & \text{for } 0 \leq t \leq 4T \\
\sin(\pi - \pi \int_0^{4T} g(t')dt'), & \text{for } 4T < t \leq 8T \\
0, & \text{else} 
\end{cases}$$

$$g(t) = \frac{1}{2\pi} \left[ Q(2\pi \cdot 0.3 \cdot \frac{t}{\sqrt{\log_2(2)}}) - Q(2\pi \cdot 0.3 \cdot \frac{t}{\sqrt{\log_2(2)}}) \right]$$

and

$$Q(t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-\frac{\tau^2}{2}} d\tau.$$ 

$T$ is the symbol period.

The base band signal is

$$y(t') = \sum_s \delta_j \cdot c_0(t'-iT + \frac{T}{2})$$

The time reference $t' = 0$ is the start of the active part of the burst as shown in figure 3. This is also the start of the bit period of bit number 0 (the first tail bit) as defined in GSM 05.02 [2].
Appendix C
Program for on-chip inductor design
Calculation of inductance, self resonance of a square spiral

\[ m := 10^{-3} \quad M := 10^6 \quad p := 10^{-12} \quad f := 10^{-15} \]

\[ \mu := 10^{-6} \quad k := 10^3 \quad n := 10^{-9} \quad G := 10^9 \]

Permeability of vacuum \[ \mu_0 := 4 \pi \cdot 10^{-7} \quad \text{H/m} \]

Frequency

\[ f_0 = 2 \ G \quad \omega_0 := 2 \pi \cdot f_0 \quad \omega_0 = 1.257 \times 10^{10} \]

Metal

\[ R_{sq} := 35 \ \text{m} \quad \text{Resistance per square} \]

\[ t_{metal6} := 1.03 \ \mu \quad \text{Thickness} \]

Conductivity

\[ \sigma := \frac{1}{R_{sq} \cdot t_{metal6}} \quad \sigma = 2.774 \times 10^7 \]

Number of turns

\[ N = 1.5 \]

No. of sides in inductor

\[ n_{sides} := 4 \cdot N \quad n_{sides} = 6 \]

Wire width

\[ W = 30 \mu \]

Spacing

\[ sp = 3 \mu \]
Appendix

Inner diameter
\[ d_{in} = 130 \mu \]

Outer diameter
\[ d_{out} := d_{in} + 2 \cdot [(N - 1) \cdot sp + N \cdot W] \quad d_{out} = 223 \mu \]
\[ d_{avg} := \frac{d_{in} + d_{out}}{2} \quad d_{avg} = 176.5 \mu \]

Skin length
\[ \delta := \frac{2}{\sqrt{\sigma \mu_0}} \quad \delta = 2.137 \mu \]
\[ \text{len} := n_{sides} \cdot d_{avg} \quad \text{len} = 1.059 \text{m} \]

Series resistance of inductor
\[ R_s := \frac{\text{len}}{\sigma \cdot W \cdot \delta \left( 1 - e^{-\frac{t_{metal6}}{\delta}} \right)} \quad R_s = 1.557 \]

Inductor resistance without accounting for the skin effect
\[ \frac{\text{len}}{\sigma \cdot W \cdot t_{metal6}} = 1.236 \]

Resistance increased by the factor below due to the skin effect
\[ \frac{t_{metal6}}{\delta \left( 1 - e^{-\frac{t_{metal6}}{\delta}} \right)} = 1.26 \]

\[ \varepsilon_{OX} := 3.4510 \cdot 10^{-11} \quad t_{metal6} := 9.08 \mu \]

Half the capacitance between metal 6 and poly
\[ C_{OX} := \frac{1}{2} \cdot \frac{\varepsilon_{OX}}{t_{metal6}} \cdot \text{len} \cdot W \quad C_{OX} = 60.373 \text{f} \]
Current sheet model for inductance value

Square spiral

\[ c_1 := 1.27 \quad c_2 := 2.067 \quad c_3 := 0.178 \quad c_3 := 0.125 \]

\[ \rho := \frac{N \cdot W + (N - 1) \cdot sp}{d_{avg}} \quad \rho = 0.263 \]

\[ L := \frac{\mu_0 N^2 \cdot d_{avg} \cdot c_1}{2} \left( \ln \left( \frac{c_2}{\rho} \right) + c_3 \rho + c_3 \rho^2 \right) \quad L = 0.666 \text{n} \]

Inductance given by Wheeler

\[ L_W := \frac{37.5 \mu_0 N^2 \left( \frac{d_{in}}{2} \right)^2}{22 \cdot \frac{d_{out}}{2} - 14 \cdot \frac{d_{in}}{2}} \quad L_W = 0.29 \text{n} \]

Simple (and bad) approximation

\[ L_2 := \mu_0 N^2 \left( \frac{d_{avg}}{2} \right)^2 \quad L_2 = 2.202 \times 10^{-5} \text{n} \]

Inductance of a solenoid with radius \( r \) and height \( h \)

ref. T. Lee p.55

\[ L_{soli}(r, h, N) := \frac{10 \pi \mu_0 N^2 \cdot r^2}{9 \cdot r + 10 \cdot h} \]

Minimum inductance       Estimated       Maximum inductance

\[ L_{soli} \left( \frac{d_{in}}{2}, 0, N \right) = 0.642 \text{n} \quad L_{soli} \left( \frac{d_{avg}}{2}, 0, N \right) = 0.871 \text{n} \quad L_{soli} \left( \frac{d_{out}}{2}, 0, N \right) = 1.1 \text{n} \]
Approximate self resonance frequency

\[ f_{SR} := \frac{1}{2 \pi \sqrt{L \cdot C_{OX}}} \quad f_{SR} = 25.1 \, \text{G} \]

Maximum useable frequency

\[ f_u := \frac{f_{SR}}{2.5} \quad f_u = 10.04 \, \text{G} \]

If the inductor is considered as a parallel resonator made up by the inductance \( L \), a loss caused by the series resistance \( R_s \) and a capacitance related to the capacitance from the wire to the patterned ground shield \( C_{ox} \)

Q of the inductor and the parallel resonator

\[ Q := \frac{\omega_0 \cdot L}{R_s} \quad Q = 5.375 \]

\[ L_p := L \left( \frac{Q^2 + 1}{Q^2} \right) \quad L_p = 0.689 \, \text{n} \]

\[ C_p := \frac{C_{OX}}{2} \left( \frac{Q^2}{Q^2 + 1} \right) \quad \text{Note 2 capacitors of Cox are in series} \]

\[ C_p = 29.177 \, \text{f} \]

\[ R_p := R_s \left( Q^2 + 1 \right) \quad R_p = 46.536 \]

Resonance of parallel resonator

Up till this frequency the resonator (the inductor) behaves as an inductor. After this frequency it looks as an capacitor.

\[ f_{presh} := \frac{1}{2 \pi \sqrt{L_p \cdot C_p}} \quad f_{presh} = 35.496 \, \text{G} \]
Appendix

Calculation of a square spiral inductor

Here are the design parameters set

- \( f_0 = 2 \cdot 10^9 \)
- \( N = 1.5 \)
- \( d_{\text{in}} = 130 \cdot 10^{-6} \)
- \( W = 30 \cdot 10^{-6} \)
- \( \text{sp} = 3 \cdot 10^{-6} \)

Results

- \( f_0 = 2 \text{ G} \)
- \( Q = 5.375 \)
- \( d_{\text{in}} = 130 \mu \text{m} \)
- \( d_{\text{avg}} = 176.5 \mu \text{m} \)
- \( d_{\text{out}} = 223 \mu \text{m} \)

Inductor model parameters

- \( L = 0.666 \text{ n} \)
- \( R_s = 1.557 \)
- \( C_{\text{OX}} = 60.373 \text{ f} \)
- \( f_{\text{SR}} = 25.1 \text{ G} \)
- \( f_u = 10.04 \text{ G} \)

Equivalent parallel resonator parameters

- \( L_p = 0.689 \text{ n} \)
- \( R_p = 46.536 \)
- \( C_p = 29.177 \text{ f} \)
- \( f_{\text{pres}} = 35.496 \text{ G} \)
Appendix D
Publications and patents

Below the published papers and patents are listed. A copy of the papers is enclosed. The papers originates from the proceeding or the official CD-Rom unless otherwise stated.

Articles on power amplifiers and linearisation:


Carsten Fallesen, Per Asbeck Nielsen, " A Highly Integrated 1 W CMOS Power Amplifier for GSM-1800 with 45% PAE", IEEE 18th Norchip conference, Turku, Finland, 6-7 November, 2000


Patents filed:
Patent filed in Great Britain, GB 0030693.6
Patent filed in Great Britain, GB 0030684.5
A RF POWER AMPLIFIER IN A DIGITAL CMOS PROCESS

Per Asbeck¹,², Carsten Fallesen¹,²
¹Nokia Denmark A/S
²Technical University of Denmark, Centre for Integrated Electronics
E-mail: Per.Asbeck@nokia.com, Carsten.Fallesen@nokia.com
Tel. +45 33 29 20 00, Fax: +45 33 29 20 01

ABSTRACT

A two stage class B power amplifier for 1.9 GHz is presented. The amplifier is fabricated in a standard digital EPI-CMOS process with low resistivity substrate. The output power is 29dBm in a 50 Ohm load. A design method, based on sweeping the loss and the resonant frequency of a LC tank to determine large signal parameters of the output transistor, is presented. Based on this method, proper values for on-chip interstage matching and off-chip output matching can be derived. Measurement of a fabricated chip is compared with the simulated circuit.

1 INTRODUCTION

The continuing reduction in production cost and the fast improvement of technology within personal communication systems makes it possible to extend the market and reach most people. The CMOS technology has played an important role in providing high functionality and complexity at low costs. For cheap wireless terminals it is attractive to integrate the RF front-end with the back-end signal processing to reduce assembly cost. Also if expensive RF technologies such as GaAs could be avoided in the design, costs could be reduced. For a RF Power Amplifier the problem is even more severe due to the limited voltage handling capability (breakdown). The reason that the integration has not been achieved is due to the lack of RF CMOS performance. The linearity and power efficiency seem to be lower given a certain power budget.

If the linearity problem could be solved, the front-end could be integrated, maybe at the price of lower power efficiency.

This is the motivation for the present work. We want to design a power amplifier that supports linear modulations schemes such as QPSK either by having sufficient linearity, or by using linearization techniques, such as Cartesian Modulation feed back, to enhance the linearity. The foundation for being able to utilize linearization techniques is that the PA itself is fairly linear to ensure stability.

In this paper we describe a class B power amplifier (PA) implemented in a mainstream digital CMOS technology. Unlike class E amplifiers, which have got a lot of attention lately [4] due to a good power efficiency, class B amplifiers are inherently more linear.
This text is divided into the following sections: section 2 discusses limitations of CMOS power amplifiers. Section 3 describes the power transistor and the matching network. Section 4 describes the gain stage, the interstage matching network and passives. Section 5 deals with the layout. Section 6 presents measurements of the fabricated chip and section 7 draws the final conclusions.

2 LIMITATIONS OF A CMOS PROCESS

The power amplifier output stage is limited by the following factors. The first important factor is the limitation in a CMOS process of a low gate oxide break down voltage (~5V for 5nm gate oxide). This voltage limits the maximum gate-drain voltage and thereby the maximum supply voltage. Depending on the class of operation (B,C,E) of the output stage, the relation between the maximum drain voltage and the supply voltage varies. For class B operation the ratio is 2, whereas for class E the ratio is 2-5 [1]. Therefore, a class E amplifier requires a lower supply voltage for the same breakdown voltage.

The second limitation is the low output impedance of the PA. To achieve a high output power, a low effective resistive load (\( R_{\text{effective}} \) in figure 1) must be chosen. The result is a high sensitivity to parasitic resistance in the matching circuits and a complicated matching circuit due to a higher transformation ratio.

The third limitation of the output stage is the maximum PA output current needed to achieve the required power in the load. The current may be so high that electromigration and parasitics in the circuit cause performance degradation.

![Figure 1. Power amplifier and load](image)

3 DESIGNING THE OUTPUT STAGE

Due to linearity and the ability to operate at low supply voltages a class B output stage is chosen. A method to find the optimum load and the large-signal effective drain-source capacitance is described.

![Figure 2. Determination of effective transistor load using a high Q tank.](image)

The size of the output transistor is determined using the circuit shown in figure 2. A high Q tank is used as a load which eliminates the effect of the drain-source capacitance. Furthermore the tank filters the harmonic of the carrier frequency so a load for the carrier frequency is found. A simulator which takes large signal nonlinearity into account is used. Harmonic balance or transient simulation are suitable. A bias voltage corresponding to the threshold voltage of the transistor is required for class B operation. The RF voltage swing should be chosen
large enough to allow the transistor to get close to the triode region. This can be determined from a standard Ids-Vgs plot.

Sweeping the resistor, $R_p$, representing the LC tank-loss, the optimum large signal working condition can be found. Such a sweep is depicted in figure 3. By varying the transistor width and sweeping the tank loss, the load for optimal output power or power efficiency can be found.

For an output power of 1 Watt and a width of 10000μm a drain efficiency of 68% was achieved. The efficiency should of course be taken with caution as losses in the matching network are neglected. The amplitude of the gate signal was 1 Volt which can easily be provided by a driver stage, as the supply voltage is 3 Volt.

![Figure 3. Simulated normalized output power (square) and PAE (cross) versus tank loss](image)

If the Q value of the tank is lowered (5-10) so the influence of the drain-source capacitance can not be neglected, a displacement in frequency of the optimum load is observed. Then the frequency displacement can be used to calculate the large signal drain-source capacitance the following way:

$$C_{ds,\text{effective}} = 2 \cdot \frac{Q}{R} \cdot \frac{\omega_{\text{peakt}}}{\omega_c} - 1$$  \hspace{1cm} (EQ 1)

Now a large signal model for the output transistor is available in the sense that an optimum resistive load and an effective drain-source capacitance are available. If the matching network establishes the effective resistance and cancels the calculated effective drain-source capacitance, the required RF operation of the transistor is ensured. The design of a matching circuit for the 50 Ohm load is now a simple matching exercise.

**4 GAIN STAGE AND INTERSTAGE MATCHING**

In this design a standard digital 0.25μm EPI-CMOS process is used. This process provides sufficient gain to achieve the desired gain of 20 dB with two stages (figure 4).

![Figure 4. Fabricated PA circuit](image)

In order to control the DC bias voltage of the output transistor independently of the drain voltage of the input transistor, a big capacitor, $C$, of 20pF is inserted. The parasitic capacitance to the substrate will be significant, because the actual capacitor requires a large area as no thin
oxide for the MIM capacitors is available. The on-chip inductor, \( L_1 \), resonates out the parasitic capacitance consisting of the gate capacitance of the output transistor and the parasitics of the on-chip DC block capacitor, \( C_p \). A simple but realistic inductor model derived from the current sheet model [6] was used in the design phase, which made it possible to find the inductor within a few iterations. Below the inductor, a grounded poly shield [7] (see chip photo figure 7) was placed to prevent capacitive coupling to the substrate. Losses will still be introduced in the low resistivity substrate (10m\( \Omega \)-cm) due to the magnetic field, which together with the resistive wire-loss limits the Q-values to a maximum of 5 or 6. The final inductor was verified by the inductor simulator program, ASITIC, and showed agreement within 5% of the simple model.

The input transistor operates in class A and is sized to cancel the losses of the inductor and to provide a voltage swing of 1V on the gate of the output transistor (see section "Designing the input stage"). The bias voltage is provided though a 400 Ohm on-chip resistor to avoid loading the AC signal on the gate.

5 LAYOUT

The layout (figure 5) of the PA was in 0.25\( \mu \)m, 6 metal layers digital CMOS process. Two large capacitors are used. The DC block capacitor in the middle of the layout was realized as lateral flux capacitors in metal 3 and metal 4, whereas metal 5 and 6 were layed out as plates on top.

![Figure 5. Layout of the PA](image)

The reason why lateral flux capacitors were used for metal 3 and 4 is that the spacing between adjacent metal is smaller (0.4\( \mu \)m) than the spacing between two different metal layers (0.9\( \mu \)m). Utilizing metal layers 3, 4, 5 and 6, a capacitance density of approximately 0.19 fF/\( \mu \)m\(^2\) is achieved. Metal layer 1 and 2 are not used, to limit the coupling to the substrate. The DC block capacitor is approximately 20pF. The grounding capacitor, \( C_{\text{gnd}} \), in figure 4 is realized using all metal layers. The density of this capacitor is approximately 0.39 fF/\( \mu \)m\(^2\). The total capacitance is 46 pF. The capacitor value is estimated based on the area of adjacent metal layers and the plate capacitor formula. The output transistor is folded to minimize the drain and source area. The maximum width of each folded transistor is 12\( \mu \)m to ensure a small gate resistance and thereby a fast transistor. Also after every 5th transistor, a column of substrate contacts is inserted to ensure a low resistance to the substrate. The ground bonding inductance should be minimized using many ground pads and bonding wires. In this design 19 ground pads were used.

The on-chip inductor was realized in metal layer 6 with a poly shield to reduce capacitive coupling to the substrate (see figure 7)
6 SIMULATIONS AND MEASUREMENT

A printed circuit board (PCB) was designed and the bare PA-chip was glued directly on the print and bonded (chip on board). Thereby the shortest possible bondwires were achieved. The maximum output power was measured as 29.2 dBm (see table 1) at the frequency of 1.95 GHz, which to our knowledge is the best performance reported for a class B power amplifier in standard digital CMOS process. In [3] and [2] an output of 13dBm and 19.3dBm respectively were achieved for a class AB PA. In [9] an output power of 31.2dBm was achieved with a CMOS process with high resistive substrate (10-20 Ω-cm) and analog options. High resistive substrate is an advantage for RF-circuits because the inductors achieve 50-100% higher Q-values and the PA-gain is higher dissipating the same amount of power. Further the process benefits from thin oxide MIM capacitors which leads to lower parasitic capacitance of the on-chip capacitors. The results achieved for the CMOS PAs [4],[5],[8] are 30dBm but these operate in class C and E which unlike class B are non-linear modes of operation. The power added efficiency of our PA was 27.4% which is approximately the same as for the two other class B CMOS amplifiers (23% in [3] and 30% in [2]).

The small signal power gain was 20.7 dB. A measured power sweep depicting the gain, PAE and output power is shown in figure 6. The used frequency is the center of the uplink frequency band of UMTS which was the design target.

The initially simulated output power was 1W and the maximum measured output power was 0.84W. The difference is due to PCB parasitics such as wire inductance and wire capacitance. Furthermore the discrete components in the matching network cause losses. We model the PCB wires with strip lines and the components in the matching network with finite Q models and achieved a result within 5% of the measured output power.

![Figure 6. Measured PA characteristic](image)

![Figure 7. Photo of entire chip and of the inductor](image)
7 CONCLUSION

A 1.9GHz class B power amplifier was designed in digital 0.25μm CMOS process. The designed and fabricated PA delivered an output power of 29.2 dBm (0.84W) which to our knowledge is the best result for a class B power amplifier in a standard digital CMOS process with low resistivity substrate. Agreement within 5% between measurement and simulation was achieved by carefully modeling wires and discrete components of the PCB. A design method based on deriving large signal parameters of the output transistor using an LC tank was presented. The design method and the models of the passives, on-chip inductor and lateral flux capacitors were described.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process</td>
<td>0.25μ CMOS</td>
</tr>
<tr>
<td>Supply voltage [V]</td>
<td>3</td>
</tr>
<tr>
<td>Frequency [GHz]</td>
<td>1.95</td>
</tr>
<tr>
<td>Max Power [dBm]</td>
<td>29.2</td>
</tr>
<tr>
<td>PAE [%]</td>
<td>27.4</td>
</tr>
<tr>
<td>small signal gain [dB]</td>
<td>20.7</td>
</tr>
<tr>
<td>large signal gain [dB]</td>
<td>14.2</td>
</tr>
</tbody>
</table>

8 REFERENCES

A Polar Linearisation System for RF Power Amplifiers

Per Asbeck¹,², Carsten Fallesen¹,²
¹Nokia Denmark A/S, ²Technical University of Denmark
E-mail: Per.Asbeck@Nokia.com

Abstract: An new transmitter architecture for polar linearization of RF power amplifiers is presented. The architecture is adapted to digital transmitters which makes it possible to move some of the analog signal processing to the digital domain. An improvement of the varying loop-gain problem in feedback linearisation systems, that is easily incorporated in the new architecture is presented. This improvement increases the stability margin and reduces the variations of the bandwidth which leads to higher linearity of the system. Simulations documenting the improvements and showing important design constraints are presented.

1. Introduction

The enormous growth in mobile phones and wireless terminals together with the limited allocation of frequency bands leads to a wish for better utilization of the bandwidth available.

The improvement needed to allow higher throughputs requires a modulation scheme that modulates the amplitude and the phase of the carrier considerably. At the transmitter side this leads to high linearity requirements to prevent a degradation of the signal quality and to prevent disturbance of neighbour channels (spectral leakage). To accommodate linearity and high power efficiency, which are normally trade offs, linearisation systems come into use. Based on predistortion, feed-forward and feedback techniques, the system tries to compensate for the non-linearity of the RF power amplifier [1].

In the following we present a new architecture for a polar linearization which uses the existing digital signal processor (DSP) in the transmitter to eliminate some analog signal processing blocks. Furthermore an enhancement to the architecture which is easily implemented in the new architecture is presented. The enhancement alleviates the loop-gain problems and the stability problems. The architectures are suited for monolithic integration [2].

System simulations based on the GSM/EDGE system [5] and a measured power amplifier (PA) are carried out to evaluate the achievable performance.

2. Polar linearisation in a digital transmitter.

A polar modulation feedback linearisation system ([2],[3]) is shown in figure 1. The system is not optimal for modern digital transceivers because the essential signals (amplitude and phase) are derived by analog circuits.

![Fig. 1. Traditional polar modulation feedback.](image_url)

We suggest an architecture (figure 2) that generates the essential signals digitally and allow us to eliminate a limiting amplifier and a mixer. The signals used for linearisation in figure 2 propagates through less analog circuits and are therefore more linear and less noisy. Furthermore, this kind of signal processing is

0-7803-6542-9/00/$10.00 © 2000 IEEE
handled more power efficient and precise in the DSP.

![Fig. 2. Polar linearisation for a digital transmitter](image)

The functionality is as follows. The DSP generates an envelope signal, \( a(t) \) and two quadrature signals, \( I'(t) \) and \( Q'(t) \), representing a RF constant envelope signal. The quadrature signals are up-converted and used as reference signal for the phase detector. The necessary extra DSP calculations are as follows:

\[
\begin{align*}
a(t) &= \sqrt{I(t)^2 + Q(t)^2} \\
I'(t) &= \frac{I(t)}{A(t)} \\
Q'(t) &= \frac{Q(t)}{A(t)}
\end{align*}
\]

where \( I(t) \) and \( Q(t) \) are the digital modulated signals normally provided by the DSP. The calculations can easily be carried by a modern DSP, if necessary by using table look up techniques. The resulting output of the system figure 2 is as follows:

\[ s(t) = A(t)(Re([I'(t) + JQ'(t)]e^{-J\omega_c t})) \] (EQ 3)

which is the normal linear modulated signal. In the following the new system will be analyzed.

3. Analysis of the polar linearisation system.

The analysis of the transmitter architecture in figure 2 is based on the signal diagram shown in figure 3. The PA is modelled by a amplitude (AM-AM) and phase (AM-PM) characteristic. The former describes the output amplitude as a function of the control voltage. The latter describes the phase difference between the RF input and output of the PA as a function of the control voltage. The loop dynamics in figure 3 is set by the loop filters \( F(s) \) and \( G(s) \). For a fixed PA-gain the amplitude and phase feed back loops are described with the following equations.

\[
\begin{align*}
\frac{a_{PA}(s)}{a(s)} &= \frac{G_{PA}G_F(s)}{1 + BG_{PA}G_F(s)} = H_{AL}(s) \\
\Phi_{PA}(s) &= H_{Ref}(s) \cdot \Phi_{Ref}(s) + H_{PM}(s) \cdot \Phi_{AMPM}(s)
\end{align*}
\]

where \( G_{PA} \) is the PA-gain, \( B \) is the gain of the amplitude detector, \( \Phi_{Ref}(s) \) is reference phase provided by the DSP and \( H_{PM}(s) \) is the injected phase distortion due to amplitude to phase conversion (AM-PM).

![Fig. 3. Polar loop signal model](image)

An overall linear system is achieved if \( H_{AL}(s) \) and \( H_{Ref}(s) \) are close to one and \( H_{PM}(s) \) is close to zero in the band of interest \( f < f_{lin} \). This bandwidth should be chosen to 2-4 times the channel bandwidth, \( f_{ch} \), to be able to attenuated spectral regrowth in the neighbour channels [1].

If we assume that a charge pump is employed within the phase loop, \( H_{Ref}(s) \) assumes an ordinary type 2 PLL transfer function [4] and \( H_{PM}(s) \) assumes a VCO phase transfer function.

The main problem of using feedback in linearisation systems is that the loop-gain decreases for a decreasing PA-gain. For our PA (figure 4) the PA-gain approaches zero for small and large amplitudes due to compression and turn-off phenomenons. Therefore, the system can only be designed to operate properly in an amplitude interval, where a minimum PA-gain is insured. For a higher order loop filter, \( G(s) \), the maximum filter gain is determined by a minimum stability gain margin, and by the maximum PA-gain, \( G_{PA,max} \). Therefore the maximum amplitude error \( (H_{AL}(s) - 1) \) cannot be minimized by choosing an arbitrarily high
filter DC gain, $G(0)$. Another performance limitation of the linearisation system is the delay between the phase, $\Phi_{PA}(t)$, and the amplitude signal $a_{PA}(t)$. The delay of the signals is determined by the closed loop transfer functions $H_{AL}(s)$ and $H_{Re}(s)$.

![Fig. 4. PA output versus the control voltage](image)

A constant delay difference between these signals could be compensated in the DSP, but a varying delay would require significant DSP processing power. From EQ. 4 we see that the amplitude transfer function varies due to varying PA-gain. The phase transfer function does not depend on the PA-gain and represents a fully linear system.

The bandwidth of amplitude transfer function varies between the bandwidth of the loop filter, $f_g$, and a bandwidth approximately $L$ times larger. Where $L$ is the loop-gain. To insure a low variation of the group delay a minimum bandwidth of the closed loop system, $H_{AL}(s)$, must be guaranteed. E.g. if a first order loop filter, $G(s)$, with one pole ($f_g$) and a DC gain of $G_0$ is used, the closed loop system is of first order, with a minimum bandwidth depending on the filter pole and the minimum PA-gain.

$$f_{CL, min} = (1 + BG_{PA, min}G_0)f_g$$  \hspace{1cm} (EQ 6)

The group delay is as follows

$$\tau(f, f_{CL}) = f_{CL}/(2\pi f^2 + f_{CL}^2)$$  \hspace{1cm} (EQ 7)

The group delay is nearly frequency independent if the closed loop bandwidth, $f_{CL}$, is significantly larger than the band of interest ($f_{lin}$).

$$\tau(f) \equiv \tau_{max} = 1/(2\pi f_{CL, min})$$  \hspace{1cm} (EQ 8)

A constant group delay results in a delayed but undistorted amplitude signal. The RF signal will, however, be distorted because of the time offset between the amplitude and phase signal. The amount of distortion also depend on the type of modulation. To relate the delay and the distortion, we calculated the distortion as a function of the delay for a long GSM/EDGE modulated [5] signal (figure 5).

![Fig. 5. ACP versus amplitude delay](image)

We found that the ACP (spectral leakage) requirements were more restrictive than the EVM (inband signal quality) requirements (EVM<9%). In figure 5 a graph is showing ACP at an offset of 200kHz and 400kHz. They are required to be below -30dB and -60dB [5] respectively. So the maximum allowable delay is 23.5ns corresponding to an ACP of -60dB at an offset of 400kHz. Based on a minimum PA-gain, EQ. 6 and EQ. 8 and the relation in figure 5, an estimate for the minimum allowable filter bandwidth can be found.

4. Gain compensation

In the previous section we showed that the bandwidth and loop-gain is a function of the PA-gain which again is a function of the amplitude of the modulated signal. For higher order systems, it’s not possible to chose the loop-gain arbitrarily high due to stability reasons. It is necessary to trade off bandwidth and loop-gain for gain stability margin.

To compensate for the varying gain in amplitude loop, we suggest to insert a variable gain stage after the calculation of the error, $e(t)$ (see figure 6). Based on the PA-gain and the amplitude, the gain of the variable gain stage can be controlled to stabilize the loop-gain.

The gain compensation signal can be realized either as an two quadrant Gilbert multiplier or as an op-amp gain stage where the feedback resistance consist of an array of switched resistors. In the former case a simple D/A converter would be required. In the latter case the gc(t) would be a n-bit bus controlling the switches.
In both cases high precision of the gain value is not required as long as the loop-gain is kept high. This means that the linearisation system is insensitive to deviations of the PA-gain e.g. due to aging and temperature variations. The gain compensation can also be non-linear without affecting the system linearity.

**Fig. 6. Linearisation system DSP controlled gain**

The gain control signal, $g_c(t)$, could be based on a sample measurement of one PA because only an approximate value is required for the compensation.

**Fig. 7. RMS amplitude error w. and w.o. gain compensation versus maximum amplitude level.**

To insure a low complexity of the D/A converter between the DSP and the multiplier the gain control signal should be quantized. The amplitude is mapped to intervals with an associated gain compensation value. The intervals and the gain control values should be chosen, so that the PA-gain times the gain control value has a minimum variation for all amplitudes. The actual gain values can be stored in a look-up table inside the DSP. The signal used for look-up is the amplitude which is already available in the DSP.

To demonstrate the principle, we simulated an amplitude loop with and without a gain control. We used the PA specified in figure 4. The gain control values were quantized to 4 bits with a minimum value of 1 and maximum value of 5. The maximum DC loop-gain was 20 and the filter pole was set to 1 MHz for both systems. The RMS amplitude error versus the maximum amplitude is shown in figure 7. An improvement is observed for all amplitude levels even though both systems have the same maximum loop-gain and stability margin. For low amplitudes we observe an improvement on more than 6% RMS.

**5. Conclusion**

A new architecture for polar linearisation of RF power amplifiers suitable for digital transmitters was presented. The number of power consuming analog signal processing blocks were reduced and moved to the digital domain (DSP). Furthermore, an extension to the new architecture that reduces the stability constraints and maintains a constant loop-gain and system bandwidth was introduced. Simulations showed the reduction in the amplitude error was more 6% for low amplitude levels.

The distortion sensitivity to delay difference in the signal path of the phase and amplitude was analyzed, and simulations showing the relation between delay and ACP were shown. Guidelines for the loop bandwidth was given. The simulations based on a radio system (GSM/EDGE) that supports amplitude and phase modulation were presented. Signal quality (EVM) and spectral regrowth (ACP) were simulated based on a measured power amplifier.

**6. References**


A 29dBm 1.9GHz Class B Power Amplifier in a digital CMOS Process

Per Asbeck1,2, Carsten Fallesen1,2
1Nokia Denmark A/S, 2Technical University of Denmark
E-mail: Per.Asbeck@Nokia.com

Abstract: A two stage class B power amplifier for 1.9 GHz is presented. The amplifier is fabricated in a standard digital CMOS process with low resistivity substrate. The output power is 29dBm in a 50 Ohm load. A design method, based on sweeping the loss and the resonance frequency of a LC tank to determine large signal parameters of the output transistor, is presented. Based on this method proper values for on-chip interstage matching and off-chip output matching can be derived. Measurement of a fabricated chip is compared with the simulated circuit.

1. Introduction

The continuing reduction in production cost and the fast improvement of technology within personal communication systems makes it possible to extend the market and reach most people. The CMOS technology has played an important role in providing high functionality and complexity at low costs. For cheap wireless terminals it is attractive to integrate the RF front-end with the back-end signal processing to reduce assembly cost. Also if expensive RF technologies such as GaAs could be avoided in the design, costs could be reduced. For RF power amplifiers the problem is even more severe due to the limited voltage handling capability (breakdown). The reason that the integration has not been achieved is due to the lack of RF CMOS performance. The linearity and power efficiency seems to be lower given a certain power budget.

If the linearity problem could be solved, the front-ends could be integrated, maybe at the price of a lower power efficiency.

This is the motivation for the present work. We want to design a power amplifier that supports linear modulations schemes such as QPSK either by possessing the linearity, or by accommodating linearization techniques, such as Cartesian Modulation feed back, to enhance the linearity. The foundation for being able to utilize linearization techniques is that the PA itself is fairly linear to insure stability.

In this paper we describe a class B power amplifier (PA) implemented in a mainstream digital CMOS technology. Opposed to class E amplifiers, which have got a lot of attention lately [2] due to a good power efficiency, class B amplifiers are inherently more linear.

This text is divided into the following chapters. Chapter 2 discusses limitations of CMOS power amplifiers. Chapter 3 describes the power transistor and the matching network. Chapter 4 describes the gain stage, the interstage matching network and passives. Chapter 5 deals with the layout. Chapter 6 presents measurements of the fabricated chip and chapter 7 draws the final conclusions.

2. Limitations in a CMOS process

The power amplifier output stage is limited by the following factors.

The first important factor is the limitation in the CMOS processes due to the low gate oxide break down voltage (~5V for 5nm gate oxide). This voltage limits the maximum gate-drain voltage and thereby the maximum supply voltage. Depending on the class of operation (B,C,E) of the output stage, the relation between the maximum drain voltage and the supply voltage varies. For class B operation the ratio is 2, whereas for class E the ratio is 2-5.
Therefore, a class E amplifier requires a lower supply voltage for the same breakdown voltage.

The second limitation is the low output impedance of the PA. To achieve a high output power, a low effective resistive load must be chosen. The result is a high sensitivity to parasitic resistance in the matching circuits and a complicated matching circuit due to a higher transformation ratio.

The third limitation of the output stage is the maximum PA output current needed to achieve the required power in the load. The current may be so high that electromigration and parasitics in the circuit cause performance degradation.

**Fig. 1. Power amplifier and load**

3. Designing the output stage

Due to the linearity and the capability to operate at low supply voltages a class B output stage is chosen. A method to find the optimum load and the large-signal effective drain-source capacitance is described.

**Fig. 2. Determination of effective transistor load using a high Q tank.**

The size of the output transistor is determined using the circuit shown in figure 2. A high Q tank is used as load which eliminates the effect of the drain-source capacitance. Furthermore the tank filters the harmonic of the carrier frequency so a load for the carrier frequency is found. A simulator which takes large signal nonlinearity into account is used. Harmonic balance or transient simulation are suited. A bias voltage corresponding to the threshold voltage of the transistor is required for class B operation. The RF voltage swing should be chosen large enough to allow the transistor to get close to the triode region. This can be determined from a standard Ids-Vgs plot.

Sweeping the LC tank-loss, the optimum large signal working condition can be found. Such a sweep is depicted in figure 3. By varying the transistor width and sweeping the tank loss, the load for optimal output power or power efficiency is found.

For an output power of 1 Watt, a width of 10000µm and a drain efficiency of 68% was achieved. The efficiency should of course be taken with cautions as losses in the matching network are neglected. The amplitude of the gate signal was 1 Volt which can easily be provided by a driver stage as the supply voltage is 3 Volt.

**Fig. 3. Normalized output power (square) and PAE (cross) versus tank loss**

If the Q value of the tank is lowered (5-10) so the influence of the drain-source capacitance can not be neglected, a displacement in frequency of the optimum load is observed. Then the frequency displacement can be used to calculate the large signal drain-source capacitance the following way:

\[
C_{ds,\text{effective}} \approx 2 \cdot \frac{Q}{\omega_c} \cdot \left(\frac{\omega_{\text{peak}}}{\omega_c} - 1\right) \quad (\text{EQ 1})
\]

Now a large signal model for the output transistor is available in the sense that a optimum resistive load and an effective drain-source capacitance is available. If the matching network establishes the effective resistance and cancels the calculated effective drain-source capacitance, the required RF operation of the transistor is insured. The design of a matching circuit for the 50 Ohm load is now a simple matching exercise.
4. Gain stage and interstage matching

In this design a 0.25 µm CMOS process is used. This process provide sufficient gain to achieve the desired gain of 20 dB with two stages (figure 4). In order to control the DC bias voltage of the output transistor independently from the drain voltage of the input transistor, a big capacitor (20pF) is inserted. The parasitic capacitance to the substrate will be significant, because the actual capacitor requires a large area as thin oxide MIM capacitors are not available.

![Fig. 4. PA circuit](image)

The on-chip inductor resonates out the parasitic capacitance consisting of the gate capacitance of the output transistor and the parasitics of the on-chip DC block capacitor. A simple but realistic inductor model derived from the current sheet model [4] was used in the design phase, which made it possible to find the inductor within a few iterations. Below the inductor, a grounded poly shield [5] (see chip photo figure 6) was placed to prevent capacitive coupling to the low resistivity epi-substrate. The final inductor was verified by the inductor simulator program, ASITIC, and showed agreement within 5% of the simple model.

The input transistor operates in class A and is sized to cancel the losses of the inductor and to provide a voltage swing of 1V on the gate of the output transistor (section 3). The bias voltage is provided through a 400 Ohm on-chip resistor to avoid loading the AC signal on the gate.

5. Layout

The layout (figure 5) of the PA was in 0.25µm, 6 metal layers CMOS process. Two area demanding capacitors are used. The DC block capacitor in the middle of the layout was realized as lateral flux capacitors in metal 3 and metal 4, whereas metal 5 and 6 were layed out as plates on top.

![Fig. 5. Layout of the PA](image)

The reason why lateral flux capacitors were used for metal 3 and 4 is that the spacing between adjacent metal is smaller (0.4µm) than the spacing between two different metal layers (0.9µm). Utilizing metal layer 3,4,5 and 6, one achieve capacitance density of approximately 0.19 fF/µm². Metal layer 1 and 2 are not used to limit the coupling to the substrate. The DC block capacitor is approximately 20pF. The grounding capacitor, C_{gnd}, in figure 4 is realized using all metal layers. The density of this capacitor is approximately 0.39 fF/µm². The total capacitance is 46 pF. The output transistor is folded to minimize the drain and source area. The maximum width of each folded transistor is 12µm to insure a small gate resistance and thereby a fast transistor. Also after every 5th transistor, a column of substrate contacts is inserted to insure a low resistance to the substrate. The ground bonding inductance should be minimized using many ground pads and bonding wires. In this design 19 ground pads were used.

The on-chip inductor was realized in metal layer 6 with a poly shield to shield against capacitive coupling to the substrate (see chip photo in figure 6)

6. Simulations and measurement

A printed circuit board (PCB) was designed and the bare PA-chip was glued directly on the print and bonded (chip on board). Thereby the shortest possible bondwires were achieved. The
maximum output power was measured to 29.2 dBm (table 1) at the frequency of 1.95 GHz.

The small signal power gain was 20.7 dB and the power added efficiency was 27.4%. A power sweep is shown in figure 7.

![Chip photo](image)

**Fig. 6. Chip photo**

![Characteristics](image)

**Fig. 7. Measured PA characteristic**

The initially simulated output power was 1W and the maximum measured output power was 0.84W. The difference is due to PCB parasitics such as wire inductance and wire capacitance. Furthermore the discrete components in the matching network causes deviations. We model the PCB wires with strip lines and the components in the matching network with finite Q models and achieved a result within 5% of the measured output power.

7. Conclusion

A 1.9GHz class B power amplifier was designed in a digital mainstream 0.25µm CMOS process. The designed and fabricated PA delivered an output power of 29.2 dBm (0.84W) which to our knowledge is the best result for a class B CMOS power amplifier [1][2][3][7]. Agreement between measurement and simulation was achieved by carefully model wires and discrete components of the PCB. A design method based on deriving large signal parameters of the output transistor using an LC tank was presented. The design of the passives, on-chip inductor and lateral flux capacitors were described.

8. References

A Highly Integrated 1 W CMOS Power Amplifier for GSM-1800

Carsten Fallesen and Per Asbeck Nielsen
Nokia Mobile Phones and Technical University of Denmark

Until now power amplifiers for handheld wireless applications have been produced almost exclusively in GaAs technologies, with a few exceptions in LDMOS, Si BJT and SiGe HBT. A CMOS power amplifier promises higher integration as well as lower cost. A typical power amplifier for wireless communication consists of 3 dies and 15-20 passive components plus decoupling. The CMOS power amplifier component count can be reduced to one die and 3-5 passives plus decoupling. This reduction in component count leads to a dramatic decrease of the cost.

The power amplifier presented in this work is targeted towards the GSM-1800 standard, which has a transmit frequency range of 1710 to 1785 MHz. The power amplifier is designed for a 0.35 µm CMOS process with 5 metal layers and metal-metal capacitors. The power amplifier consists of two stages with integrated input and interstage impedance matching networks as well as the very first part of the output matching network. The two stages of the power amplifier both operate in class AB. The schematics of the power amplifier are shown in Figure 1 and Figure 2.

The input matching network is a fully integrated highpass LC matching section. This was chosen because it incorporates DC blocking and biasing at the same time as the matching. The inductor is a spiral inductor implemented in top metal layer, while the capacitor is made with the two lowest metal layers. The transistor at the input stage is 1 mm wide and 0.35 µm long. The input stage has an off-chip inductor acting as a RF choke (RFC). The interstage matching network consists of a LC highpass section for the same reasons as the input matching network. The implementation is made in the same way as the input matching. The output transistor is 8 mm wide and has a length of 0.35 µm, it is partitioned into 6 separate finger transistors, with 30 fingers each. The output matching network is placed primarily off-chip due to efficiency considerations. In order to have better harmonic termination, a capacitor is placed on-chip, directly at the drain of the transistor, in parallel with the drain-source capacitor of the transistor. This capacitor terminates the harmonics at the drain, but at the same time it transforms the output impedance even lower. The RFC for the output stage is incorporated in the output matching network. The output matching network consists of a bandpass T section, since this allows control of the Q of the otherwise high transformation factor.

The CMOS power amplifier IC has been mounted directly on the PCB, and wire bonded directly on a goldplated PCB microstrips. The passive components used on the PCB are 0402 and 0603 SMD components. The SMA connectors have been mounted horizontally on the edge of the PCB, in order to reduce the effects of the transition from SMA connector to PCB microstrip.

The measurements were made in pulsed mode according to the GSM1800 specification, with a supply voltage of 3.5 V. The highest power added efficiency was 40% at 1730 MHz, with an output power of 30.3 dBm. The output power and efficiency biased for maximum power added efficiency vs. frequency is shown in Figure 4. The highest output power obtained was 31.5 dBm at 1670 MHz, with the input power increased to 15 dBm. The output power and efficiency with biases set for maximum output power vs. frequency is shown in Figure 3. The power amplifier operates on a supply voltage from 1 V to 4 V. The output power and efficiency vs. supply voltage is shown in Figure 5. The maximum output power is 32.2 dBm at 4 V.

One of the most important aspects of the reliable design of a power amplifier is accurate simulations, and a lot of emphasis was put on simulations. The simulations were made in the APLAC simulator, primarily with small-signal and harmonic balance simulations, but transient simulations have also been used. The simulated and measured data were compared, a few minor modeling problems were found and corrected. The simulations showed very good agreement between the simulated and measured results, the output power were predicted within a few tenths of a dB. The comparison between simulated and measured results is shown in Figure 6.

A redesign with minor adjustments of the input and interstage matching networks will move maximum output power and efficiency to the GSM-1800 band. The die area including pads is 1.9 sq. mm. The power amplifier consists of one die, two RFCs and three matching component plus decoupling capacitors, compared to 3 dies and 15-20 passives plus decoupling capacitors for a typical GaAs power amplifier. This power amplifier gives the highest output power for a CMOS amplifier in the 1.8 GHz range. The efficiency is better than other CMOS amplifiers using linear modes[1] and comparable to or better than the switched mode approaches typically used in CMOS [2][3].

Figure 1  Schematic of the power amplifier.

Figure 2  Schematic of the output matching network.

Figure 3  Measured output power and efficiency, biased for maximum output power.

Figure 4  Measured output power and efficiency, biased for maximum power added efficiency.

Figure 5  Measured output power and efficiency vs. supply voltage at 1670 MHz.

Figure 6  Comparison of simulated and measured data.

Table 1. Characteristics of the power amplifier

<table>
<thead>
<tr>
<th>Process</th>
<th>0.35 µm CMOS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply voltage</td>
<td>3.5 V</td>
</tr>
<tr>
<td>Input power</td>
<td>5 dBm</td>
</tr>
<tr>
<td>Output power</td>
<td>31.5 dBm</td>
</tr>
<tr>
<td>Frequency</td>
<td>1710-1785 MHz</td>
</tr>
<tr>
<td>Max. drain efficiency</td>
<td>45%</td>
</tr>
<tr>
<td>Max. total efficiency</td>
<td>40%</td>
</tr>
<tr>
<td>Die area</td>
<td>1.9 sq. mm</td>
</tr>
</tbody>
</table>

Figure 7  Photograph of the power amplifier IC
A Highly Integrated 1W CMOS Power Amplifier for GSM-1800 with 45% PAE

Carsten Fallesen and Per Asbeck
Nokia Denmark A/S and Technical University of Denmark
Phone: +45 33292467  Fax: +45 33292001
Carsten.Fallesen@nokia.com

This paper presents a power amplifier designed in a 0.35 μm CMOS process. A CMOS power amplifier promises the possibility of lower cost and a higher degree of transmitter integration. The input and interstage matching networks have been fully integrated, thereby reducing the number of external components to three plus decoupling capacitors. The key features of the power amplifier are class AB operation with 31.2 dBm output power at 1730 MHz, and a maximum power added efficiency of 45%. The nominal supply voltage is 3.5 V, but the power amplifier operates down to 1.0 V. The efficiency and output power is better than reported for other CMOS amplifiers whether they use linear modes of operation or the switched mode approaches typically used in CMOS.

1. INTRODUCTION

Until now power amplifiers for wireless applications have been produced almost exclusively in GaAs technologies, with a few exceptions in LDMOS, Si BJT and SiGe HBT. The submicron CMOS processes are now usable for power amplifier design, and are without doubt the cheapest processes available. Due to the high yield in CMOS fabrication, higher integration is possible. A CMOS power amplifier therefore promises both higher integration and lower cost. A typical power amplifier module for wireless communication consists of 3 dice and 15-20 passive components plus decoupling. The CMOS power amplifier component count can be reduced to one die and 3-5 passives plus decoupling. This reduction in component count leads to a significant reduction in power amplifier cost.

Until recently, linearity of power amplifiers have not been a problem in most wireless standards. This was due to the fact that most systems, such as GSM, were constant envelope modulated, meaning that no information was stored in the amplitude. But non-constant envelope systems, such as IS-95 and WCDMA, have introduced the need for linear power amplifiers.

The problem with nonlinear power amplifiers and amplitude modulated systems is caused by spectral regrowth, due to the AM-PM conversion in the power amplifier. This means that the modulated signal will leak into the neighboring channels. The leakage is characterized by the adjacent channel power ratio (ACPR), relating the power in the channel to the power leaked into the neighboring channel.

The requirement of IS-95 is an ACPR of 26 dB. This is, however, only the start, in the 3G WCDMA wireless standards the requirement is 42 dB. Because of this, more effort will have to be placed in the design of linear power amplifiers. The linearity can be achieved by designing class A amplifiers with low efficiency or by applying linearization techniques to relatively linear power amplifiers.

The power amplifier presented in this work is targeted towards the GSM-1800 standard, which has a transmit frequency for the handset of 1710 to 1785 MHz. The goal has been to design a power amplifier with a 1 W output power with a linearity sufficient to handle the transition to 3G mobile standards, including the upcoming EDGE standard. The linearity will be improved using linearization techniques in either software or hardware. However, it is
important that the starting point is reasonable, this is the case for the designed class AB power amplifier.

2. THE DESIGNED CIRCUIT

![Schematic of the power amplifier.](image)

The power amplifier is designed for a 0.35 μm bulk CMOS process with a substrate resistivity of 10-20 Ω-cm. The process has 5 metal layers and thin-oxide metal-metal capacitors. The special metal-metal capacitors have a high density and therefore the die size (cost) of the complete power amplifier can be reduced.

The power amplifier consists of two stages with integrated input and interstage impedance matching networks as well as the very first part of the output matching network. The schematic of the power amplifier is shown in Figure 1. The input matching network transforms the conjugate gate impedance to 50 Ω and cancels the effect of the bondwires. It is made with a fully integrated highpass LC matching section. This has been chosen because it incorporates DC blocking and biasing at the same time. The inductor is a spiral inductor implemented in the top metal layer, while the capacitor is made with the two lowest metal layers.

The input stage operates in class AB, delivering up to 15 dB gain at maximum output power. The transistor of the input stage is 1 mm wide and 0.35 μm long. The input stage has an off-chip inductor acting as a RF choke (RFC).

The interstage matching network consists of a LC highpass section for the same reasons as the input matching network. The implementation is made in the same way as the input matching, but the interstage matching transforms the gate impedance of the output transistor to the desired output load of the input stage. The output stage operates in class AB close to class B. There are a number of reasons to choose this mode of operation:

1. Class AB close to class B is relatively linear. This is not the case for class C and E amplifiers. The linearity is, however, not as good as class A.
2. The efficiency is relatively good, the theoretical maximum is 78.5%, compared with 50% for the class A amplifiers and Class C and E amplifiers have theoretical efficiencies of up to 100%.
3. The maximum drain voltage is twice the supply voltage, this is important due to the possible breakdown of the gate-oxide. Class C and E amplifiers easily exceed three times the supply voltage.
4. The power utilization factor (PUF), which is a measure of the gain compared to the output power, is reasonable compared with class A, and better than class C and E.
5. The required output load impedance is not too low to implement efficiently, which is often the case for class C.

The output transistor is 8 mm wide and has a length of 0.35 μm as the input stage. The transistor is partitioned into 6 separate finger transistors, with 70 fingers each. The gain of the output stage is approximately 12 dB.

The output matching network is placed primarily off-chip due to efficiency considerations. In order to have better harmonic termination, a capacitor is placed on-chip, directly at the drain of the transistor, in parallel with the drain-source capacitor of the transistor. This capacitor terminates the harmonics at the drain, but at the same time it transforms the output impedance even lower, leaving a more difficult matching problem. The RFC for the output stage is a relatively short microstrip which can be implemented without increasing the PCB size. The output matching network consists of a bandpass T section, due to the high transformation factor from 4 Ω to 50 Ω. The choice of the T section gives a larger bandwidth than a single L section. In the design of the network, the parasitics of the RFC microstrip has to be included.

In the design phase the two amplifier stages were initially treated separately and optimized for 50 Ω input and output matching using load-pull simulations. After each stage had been optimized an interstage matching network was designed using the impedances found for each stage.

3. SIMULATIONS

In order to make precise simulations the PCB was characterized using simple short, open and through structures. Using this approach, the transients in the SMA connector to PCB interface could be modeled accurately. The simulations of these test structures were very precise up to 4 GHz. After the SMA connectors and PCB were characterized, the output matching network was simulated and measured.

The PCB has been modeled with microstrip lines and the SMD components have been modeled according to vendor specifications. The bondwires have been modeled as inductors and the mutual coupling between the bondwires were included as well.

The measurements of the output matching network, were performed by mounting a short piece of semi-rigid cable in place of the IC. The shield of the cable was soldered to the ground plane, where the IC was supposed to be placed. The conductor of the cable was attached at the microstrip on the PCB where the bondwires from the output of the IC would go.

The IC simulations were based on parasitic extraction from the transistor and passive layout. The transistors were modeled by the MOS9 model. The spiral inductors and the metal-metal capacitors were simulated using lumped models delivered by the foundry.

![Graph](image)

Figure 2 Comparison of simulated and measured data.
The simulations were made in the APLAC simulator, primarily with small-signal and harmonic balance simulations. For verification purposes transient simulations have been performed in the Eldo simulator. The harmonic balance simulations proved to be faster than the transient simulations, since only the steady-state solution is calculated, the precision of the simulations proved to be the same.

The simulations for the complete power amplifier including the PCB showed very good agreement between the simulated and measured results. The measured output power was predicted within a few tenths of a dB. The comparison between simulated and measured output power is shown in Figure 2.

4. MEASUREMENTS

![Graph](image1)

Figure 3  Output power and efficiency vs. frequency, biased for maximum efficiency

![Graph](image2)

Figure 4  Output power and efficiency vs. frequency, biased for maximum output power

The CMOS power amplifier IC was been mounted directly on the PCB, and wire bonded directly onto the PCB microstrips. To enable the wire bonding, the PCB was gold plated, the dielectric used in this work was standard FR4, with a relative dielectric constant of approximately 4.3 at 1.75 GHz.

The PCB can be produced either with a PCB milling machine or at a normal PCB production facility. The first approach offers some advantages during the prototyping phase of the design. A new PCB can be built within hours, allowing for larger exploration of the design space, particularly the topology of the output matching network.

The passive components used on the PCB were 0402 and 0603 SMD components. The SMA connectors were mounted horizontally on the edge of the PCB, in order to reduce the effects of the transition from SMA connector to PCB microstrip.

To get a realistic picture of the performance, the measurements were made in pulsed mode according to the GSM1800 specifications, this means a duty cycle of 12.5%.

The highest power added efficiency was 45% at 1730 MHz, with an output power of 30.4 dBm. The output power and efficiency measurements with the power amplifier biased for maximum power added efficiency vs. frequency are shown in Figure 3.

The highest output power obtained was 31.3 dBm at 1720 MHz. The output power and efficiency measurements with biases set for maximum output power vs. frequency are shown in Figure 4.

The power amplifier operates on a supply voltage from 1 V to 4 V. The output power and efficiency vs. supply voltage is shown in Figure 5. The maximum output power is 32.2 dBm at 4 V.
Figure 5 Measured output power and efficiency vs. supply voltage at 1730 MHz.

Due to a mismatch in the input impedance matching network, the gain of the power amplifier is lower than predicted at design time. The mismatch was caused by an error in the estimation of the parasitics of input pins.

Since the power amplifier is operating in class AB, it is inherently more linear, than e.g. the class C, D and E amplifiers demonstrated in CMOS so far [1][3][5][6]. This means that the power amplifier is suitable for digital predistortion. The adjacent channel power of the power amplifier was made without any optimizations towards lower ACPR. The measurements showed that the ACPR requirements of EDGE was met up to 2 dB from maximum required output power. At the maximum required output power the ACPR was -32 dBC, whereas the requirement of EDGE is -40 dBC. It has been shown, that an improvement of the adjacent channel power ratio (ACPR) of 8-10 dB, is possible with simple low-power digital predistortion [7][8]. This means that the power amplifier can be used for EDGE if a simple digital predistortion system is incorporated into the DSP.

A comparison of all the published CMOS power amplifier results is shown in Table 1. As can be seen from the table no other CMOS power amplifier has been published with a output power or efficiency as high as the work presented here.

<table>
<thead>
<tr>
<th></th>
<th>Frequency (MHz)</th>
<th>P_out (dBm)</th>
<th>PAE (%)</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. Melly et. al. [1]</td>
<td>430</td>
<td>4.0</td>
<td>15</td>
<td>C</td>
</tr>
<tr>
<td>S.-J. Yoo et. al. [2]</td>
<td>433</td>
<td>13.0</td>
<td>30</td>
<td>AB</td>
</tr>
<tr>
<td>D. Su et. al. [3]</td>
<td>830</td>
<td>30.0</td>
<td>42</td>
<td>D</td>
</tr>
<tr>
<td>B. Ballweber et. al. [4]</td>
<td>900</td>
<td>19.3</td>
<td>23</td>
<td>AB</td>
</tr>
<tr>
<td>C. Yoo et. al. [5]</td>
<td>900</td>
<td>29.5</td>
<td>41</td>
<td>E</td>
</tr>
<tr>
<td>K.-C. Tsai et. al. [6]</td>
<td>1980</td>
<td>30.0</td>
<td>41</td>
<td>E</td>
</tr>
<tr>
<td>Asbeck et. al. [9]</td>
<td>1950</td>
<td>29.2</td>
<td>27</td>
<td>B</td>
</tr>
<tr>
<td><strong>This work</strong></td>
<td><strong>1730</strong></td>
<td><strong>30.4</strong></td>
<td><strong>45</strong></td>
<td><strong>AB</strong></td>
</tr>
</tbody>
</table>

Table 1. Comparison of CMOS Power Amplifiers
5. CONCLUSION

A CMOS power amplifier has been presented with a power added efficiency of 45% with an output power of 30.4 dBm at 1730 MHz. When biased for maximum output power 31.2 dBm is delivered while maintaining an efficiency of 42%. The die area including pads is 1.9 sq. mm. By accurately modeling bondwires, microstrips and SMD components the accuracy of the simulations was improved, and is now within a few tenths of a dB, compared to measured results.

The power amplifier consists of one die, one RFC, one microstrip and two matching components plus decoupling capacitors, compared to 3 dice and 15-20 passives plus decoupling capacitors for a typical GaAs power amplifier.

The power amplifier operates in class AB, which gives good output power, efficiency and linearity. Until now no CMOS power amplifiers with this output power or efficiency have been published. The fact that this power amplifier is relatively linear, means that it is useful in wireless applications, especially in systems which utilize amplitude modulation. With a digital predistortion system the power amplifier can be used for EDGE.

6. REFERENCES

[5] Changsik Yoo and Quiting Huang, “A Common-Gate Switched, 0.9W Class-E Power Amplifier with 41% PAE in 0.2 μm CMOS”, 2000 Symposium on VLSI Circuits, pp. 56-57.
10.3 A 1W 0.35μm CMOS Power Amplifier for GSM-1800 with 45% PAE

Carsten Fallesen, Per Asbeck
Nokia Denmark A/S, Copenhagen, Denmark

Until now, power amplifiers for wireless applications have been produced almost exclusively in GaAs technologies, with a few exceptions in LDMOS, Si BJT and SiGe HBT. The submicron CMOS processes are now usable for power amplifier design, and are without doubt the cheapest processes available and high integration is possible due to high yield. A typical power amplifier module for wireless communication consists of three die and 15-20 passive components plus decoupling. This CMOS power amplifier has one die and three passives plus decoupling. This reduction in component count leads to a significant reduction in power amplifier cost.

The power amplifier presented in this work is targeted at the GSM-1800 standard, which has handset transmit frequency of 1710MHz to 1755MHz. It is designed for a 0.35μm bulk CMOS process with a substrate resistivity of 10-20Ωcm. The process has five metal layers and thin-oxide metal-metal capacitors. The thin-oxide metal-metal capacitors are high density and therefore the die size (cost) of the complete power amplifier is reduced.

The power amplifier consists of two stages with integrated input and interstage matching. The schematic of the power amplifier is shown in Figure 10.3.1. The input matching network is made with a fully-integrated highpass LC matching section. This is chosen because it incorporates DC blocking and biasing at the same time. The inductor is a spiral inductor implemented in the top metal layer, while the capacitor is made by the two lowest metal layers.

The input stage operates in class AB, delivering up to 15dB gain. The transistor in the input stage is 1mm wide and 0.35μm long. The input stage has an off-chip inductor acting as an RF choke. The interstage-matching network consists of a LC highpass section for the same reasons as the input matching network. The implementation is made in the same way as the input matching.

The output stage operates in class AB close to class B. The output transistor is 8mm wide and has a length of 0.35μm as does the input stage. The transistor is partitioned into 6 separate transistors, with 70 fingers each. The gain in the output stage is ~12dB. The output-matching network is placed primarily off-chip due to efficiency. To improve harmonic termination, a capacitor is placed on-chip, directly at the drain of the transistor. This capacitor terminates the harmonics at the drain, but at the same time it transforms the output impedance further down, leaving a more difficult matching problem. The RF choke for the output stage is a relatively short microstrip, which can be implemented without increasing the PCB size. The output-matching network consists of a bandpass T section, due to the high transformation factor from 4Ωto 50Ω. The choice of the T section gives a larger bandwidth than a single L section.

The CMOS power amplifier IC is mounted directly on the PCB and wire bonded directly on to the PCB microstrips. To enable the wire bonding, the PCB is gold plated. The dielectric used in this work is standard FR4, with a relative dielectric constant of ~4.3 at 1.75GHz.

The highest power added efficiency is 45% at 1730MHz, with an output power of 30.4dBm. The output power and efficiency measurements of the power amplifier biased for maximum power added efficiency vs. frequency are shown in Figure 10.3.2. The highest output power obtained was 31.3dBm at 1720MHz. The output power and efficiency measurements with biases set for maximum output power vs. frequency are shown in Figure 10.3.3. This power amplifier reaches higher power added efficiency and output power than any of the CMOS power amplifier previously published [1][2][3][4]. The power amplifier operates on supply voltages as low as 1V, the measured output power and efficiency vs. supply voltage are shown in Figure 10.3.4.

Simulations of the complete power amplifier including the PCB are in good agreement with measured results. The measured output power is predicted within a few tenths of a dB. Comparison between simulated and measured output power is shown in Figure 10.3.5. Since the power amplifier operates class AB, it is inherently more linear, than e.g. the class D and E amplifiers thus far reported in CMOS [1][3][4]. This means that the power amplifier is suitable for digital predistortion.

Adjacent channel power measurements of the power amplifier are without any optimizations for lower adjacent channel power ratio (ACPR). The measurements show the ACPR requirements of EDGE standard are met up to 2dB from maximum required output power. At maximum required output power the ACPR is ~32dBc, whereas the requirement of EDGE is ~40dBc. It is shown that an improvement of the adjacent channel power ratio (ACPR) of 8-10dB is possible with simple low-power digital pre-distortion [5]. This means that the power amplifier can be used for EDGE if a simple digital predistortion system is incorporated into the DSP.

A power amplifier for GSM-1800 demonstrated in a 0.35μm CMOS process has area including pads 1.9mm². The efficiency and output power are better than those of any other CMOS power amplifiers published and the linearity of the power amplifier is sufficient for EDGE.

References:
[3] C. Yao and Q. Huang, "A Common-Gate Switched, 0.9W Class-E Power Amplifier with 41% PAE in 0.2mm CMOS", 2000 Symposium on VLSI Circuits, pp. 56-57, 2000
A 1W CMOS POWER AMPLIFIER FOR GSM-1800 WITH 55% PAE

Carsten Fallesen\textsuperscript{1,2,*} and Per Asbeck\textsuperscript{1,2}

\textsuperscript{1}Nokia Denmark, Copenhagen, DK-1790, Denmark
\textsuperscript{2}Technical University of Denmark, Lyngby, DK-2800, Denmark
\textsuperscript{*}Now with Thrane & Thrane, Lyngby, DK-2800, Denmark

Abstract -- Until recently it was the common opinion that CMOS RF power amplifiers were not feasible for mobile handsets. This paper presents a CMOS power amplifier for the GSM-1800 standard, with only two external matching components and a few decoupling capacitors. The performance of the power amplifier is better than any other CMOS power amplifier reported and comparable to commercially available power amplifiers in other technologies.

I. INTRODUCTION

This paper presents the results achieved in the design of a 1W CMOS power amplifier for GSM-1800. Due to the high yield in CMOS fabrication, higher integration is possible than e.g. in GaAs processes. A CMOS power amplifier therefore promises higher integration as well as lower cost. A typical power amplifier module for wireless communication consists of 2-3 dice and 15-20 passive components. The CMOS power amplifier component count can be reduced to one die and 2-5 passives plus decoupling. This reduction in component count leads to a significant reduction in power amplifier cost.

The power amplifier presented in this work is targeted towards the GSM-1800 standard, which has a transmit frequency for the handset of 1710 to 1785 MHz. The goal has been to design a power amplifier with a 1 W output power.

II. DESIGN

The design of this power amplifier followed the design and simulation methodologies described in [1]. The design is the second iteration of a 1W CMOS power amplifier for GSM-1800, results of the first iteration have previously been presented [2]. This power amplifier shows higher integration and much better efficiency than previously presented.

The power amplifier is designed for a 0.35 \textmu m bulk CMOS process with a substrate resistivity of 10-20 \textOmega cm. The process has 5 metal layers and thin-oxide metal-metal capacitors. The thin-oxide metal-metal capacitors have a high density and therefore the die size (cost) of the complete power amplifier can be reduced.

The first choice to make was the number of stages in the power amplifier. In this case a two-stage methodology was chosen.

Then the class of operation was chosen for each of the stages. The input and output stages operates in class AB close to class B. There are a number of reasons to choose this mode of operation:

1. Class AB close to class B is relatively linear. This is not the case for class C and E amplifiers. The linearity is, however, not as good as class A.
2. The efficiency is relatively good, the theoretical maximum is 78.5%, compared with 50% for the class A amplifiers and Class C and E amplifiers have theoretical efficiencies of up to 100%.
3. The maximum drain voltage is twice the supply voltage, this is important due to the possible breakdown of the gate-oxide. Class C and E amplifiers easily exceed three times the supply voltage.
4. The power utilization factor (PUF), which is a measure of the gain compared to the output power, is reasonable compared with class A, and better than class C and E.
5. The required output load impedance is not too low to implement efficiently, which is often the case for class C.

Once the class of operation was chosen for the output stage it was possible to start the dimensioning of the output transistor. This dimensioning was an iterative process where the initial guess originated from the I-V characteristic of the power amplifier. From the I-V characteristic it was possible to find the voltage and current swings possible for a given load-line. From the voltage and current swings the maximum output power was then determined and a reasonable size of the transistor was found.

After an initial value is selected the more accurate RF behavior is found using load-pull simulations. The load-pull simulations are the simulation equivalent of the load-pull measurements.

The final schematic of the power amplifier is shown in Fig 1 where the components mentioned below can be located. The output transistor (M\textsubscript{2}) was then chosen to be 8...
mm wide and with a length of 0.35 µm. The transistor is partitioned into 6 separate finger transistors, with 70 fingers each. The input stage also operates in class AB. The transistor of the input stage (M1) is 1 mm wide and 0.35 µm long. The load-pull simulation results for the transistor is shown in Fig 1.

The output matching network is placed primarily off-chip due to efficiency considerations. In order to have better harmonic termination, a capacitor (C1) is placed on-chip, directly at the drain of the transistor, in parallel with the drain-source capacitor of the transistor. This capacitor terminates the harmonics at the drain, but at the same time it transforms the output impedance even lower, leaving a more difficult matching problem. The RF chokes (RFC1, RFC2) for the output stage as well as the input stage are relatively short microstrips, which can be implemented without increasing the overall PCB size. The output matching network consists of a bandpass T section (L1, C2, C3), due to the high transformation factor from 4 Ω to 50 Ω. The choice of the T section gives a larger bandwidth than a single L section. The inductor in the T section consists of a contribution from the bondwires as well as from the microstrip. In the design of the network, the parasitics of the RFC microstrip were also included.

The input and interstage matching networks were both made with a fully integrated highpass LC matching section. This was chosen because it incorporates DC blocking and biasing at the same time as the impedance matching. The on-chip inductors are spiral inductors implemented in the top metal layer, while the capacitors are made thin-oxide metal-metal capacitors in the two lowest metal layers.

### III. Simulation

Although it is possible for modern CAD tools to extract every parasitic component in a layout, it is not desirable to blindly use those parasitics. In general the slowdown of the simulator will be significant if all parasitics are included. During the design phase the designer will have to make some decisions, although complexity of the simulations will increase.

There are a large number of simulation models available for MOS transistors. The gate resistance is usually not included in the transistor models from the vendor, it is therefore important to take it into account. In this work the MOS9 model has been used.

Over the last couple of years a lot of research effort has been put into the characterization and modeling of the on-chip inductors. The capacitor can be modeled as the intended capacitor, with an additional capacitor from the bottom plate to the substrate. A parasitic series resistance is associated with each of the two plates.

One of the important things when trying to simulate a complete chip is to maintain the overview. If all interconnects were modeled regardless of their impact on the performance, the simulation speed would increase drastically and hence prevent simulation of the complete chip. It is therefore important to carefully select which interconnects should be modeled.

The bondwires are an important part of the entire RF design. The bondwires have inductance, capacitance and resistance associated with them. The PCB is modeled using microstrips. In some simulator e.g. APLAC, a number of microstrip components are implemented. The models from the vendor have been sufficient to get accurate simulations. Other issues like thermal modeling of the pulsed power amplifier have also been addressed.

In the design phase a number of different simulation methods are used. The most important method is harmonic balance simulations to get the steady-state response from the circuit. Another important simulation is the classical small-signal simulation to get initial design guesses and various small-signal parameters such as small-signal gain and stability. At last the transient simulations are also used to obtain information about time-domain phenomena such as stability, modulation and power ramping.

### IV. Measurements

The CMOS power amplifier IC was been mounted directly on the PCB and wire bonded directly onto the PCB microstrips. To enable the wire bonding, the PCB was gold plated, the dielectric used in this work was standard FR4, with a relative dielectric constant of approximately 4.3 at 1.75 GHz.
The passive components used on the PCB were 0402 SMD components. The SMA connectors were mounted horizontally on the edge of the PCB, in order to reduce the effects of the transition from SMA connector to PCB microstrip.

To get a realistic picture of the performance, the measurements were made in pulsed mode according to the GSM1800 specifications, this means a duty cycle of 12.5%.

The highest power added efficiency was 55% at 1750 MHz, with an output power of 30.4 dBm. The output power and efficiency measurements with the power amplifier biased for maximum power added efficiency vs. frequency are shown in Fig 2.

The power amplifier operates on a supply voltage from 1 V to 3.4 V. The output power and efficiency vs. supply voltage is shown in Fig 3. The output power is 20.8 dBm at 1V and 31.4 dBm at 3.4 V. The power added efficiency varies from 43% to 55% at 1V and 3.4V respectively.

A comparison of all the published CMOS power amplifier results is shown in Table I. As can be seen from the table no other CMOS power amplifier has been published with output power or power added efficiency as high as the work presented here.

The simulations for the complete power amplifier including the PCB showed very good agreement between the simulated and measured results. The measured output power was predicted within a few tenths of a dB. The efficiency deviated less than 1%. The comparison between simulated and measured output power is shown in Fig 4.
V. Conclusion

A CMOS power amplifier has been presented with a power added efficiency of 55% with an output power of 30.4 dBm at 1750 MHz. The power amplifier is designed for GSM-1800 with a supply voltage of 3V, although it performs very well from 1V to 3.4V. The die area including pads is 1.1 sq. mm. By accurately modeling bondwires, microstrips and SMD components the accuracy of the simulations was within a few tenths of a dB, compared to measured results.

The power amplifier consists of one die, two short microstrips and two matching components plus decoupling capacitors, compared to 3 dice and 15-20 passives plus decoupling capacitors for a typical GaAs power amplifier. The power amplifier has higher power added efficiency than any other CMOS power amplifier results published so far, whether they operate linearly \[3\][4][5][2] or non-linearly [6][7][8][9].

References


[8] C. Yoo and Q. Huang, "A common-gate switched, 0.9W class-E power amplifier with 41% PAE in 0.2 um CMOS," in 2000 Symposium on VLSI Circuits, pp. 56--57, 2000.


Table I

<table>
<thead>
<tr>
<th></th>
<th>Frequency (MHz)</th>
<th>P_out (dBm)</th>
<th>PAE (%)</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. Melly et. al. [6]</td>
<td>430</td>
<td>4.0</td>
<td>15</td>
<td>C</td>
</tr>
<tr>
<td>S.-J. Yoo et. al. [3]</td>
<td>433</td>
<td>13.0</td>
<td>30</td>
<td>AB</td>
</tr>
<tr>
<td>D. Su et. al. [7]</td>
<td>830</td>
<td>30.0</td>
<td>42</td>
<td>D</td>
</tr>
<tr>
<td>B. Ballweber et. al. [4]</td>
<td>900</td>
<td>19.3</td>
<td>23</td>
<td>AB</td>
</tr>
<tr>
<td>C. Yoo et. al. [8]</td>
<td>900</td>
<td>29.5</td>
<td>41</td>
<td>E</td>
</tr>
<tr>
<td>K.-C. Tsai et. al. [9]</td>
<td>1980</td>
<td>30.0</td>
<td>41</td>
<td>E</td>
</tr>
<tr>
<td>Asbeck et. al. [5]</td>
<td>1950</td>
<td>29.2</td>
<td>27</td>
<td>B</td>
</tr>
<tr>
<td>Fallesen et. al. [2]</td>
<td>1730</td>
<td>30.4</td>
<td>45</td>
<td>AB</td>
</tr>
<tr>
<td>This work</td>
<td>1750</td>
<td>30.4</td>
<td>55</td>
<td>AB</td>
</tr>
</tbody>
</table>