Design considerations for RF power amplifiers demonstrated through a GSM/EDGE power amplifier module

DOI:
10.1007/0-306-47950-8_13

Document status and date:
Published: 01/01/2002

Document Version:
Accepted manuscript including changes made at the peer-review stage

Design Considerations for RF Power Amplifiers demonstrated through a GSM/EDGE Power Amplifier Module

Peter Baltus and André van Bezooijen
Philips Semiconductors
Gerstweg 2
6534 AE Nijmegen

Abstract

This paper describes the design considerations for RF power amplifiers in general, including trends in systems, linearity and efficiency, the PA environment, implementation issues and technology.

As an example a triple-band (900/1800/1900MHz) dual mode (GSM/Edge) power amplifier module is described in this article. The RF transistors and biasing circuitry are implemented in silicon bipolar technology. A multi-layer LTCC substrate is used as carrier.

1. Introduction

Currently, many cellular systems are in use in different regions of the world, and in many places more than one system is in use simultaneously. In Europe and Asia the dominant system is currently GSM, in the US it is IS95, but also AMPS and GSM-like systems co-exist, and in Japan PHS, PDC and IS95 co-exist.

The handsets for these systems use a low-power transceiver to communicate with a network of base stations using radio transmissions in the frequency range of 800MHz to 2500MHz, and transmit power levels in the range of 10mW to 2W. **Table 1** below shows an overview of important cellular systems.

	Frequency (MHz)		Power (dBm)		Modulation
	Min	*Max*	*Min*	*Max*	
AMPS	824	849	7	35	FM
PHS	1895	1918	-22	18	QPSK
GSM	880	915	5	33	GMSK
GSM/EDGE	880	915	5	27	3π/8-8PSK
DCS	1710	1785	0	30	GMSK
DCS/EDGE	1710	1785	0	26	3π/8-8PSK
PCS	1850	1910	0	30	GMSK
PCS/EDGE	1850	1910	0	26	3π/8-8PSK
IS95	824	849	-53	27	O-QPSK
UMTS	1920	1980	-44	24	HPSK

Table 1 Properties of cellular systems

1.1 Trends in cellular systems

New so-called third generation (3G) systems are being introduced which will allow for higher capacity (more users and higher user bit rates), and better compatibility across the world. Several of these systems are covered by the IMT2000 standard and include W-CDMA/UMTS (Japan and Europe) and CDMA2000 (USA).

In most countries, there will be a gradual change from existing second generation (2G) systems to future 3G systems, with both systems coexisting for at least a few years. For this reason, and because no single standard will achieve world-wide coverage in the near future, handsets that can connect to multiple systems will be required. Such handsets will provide access to the features of the 3G networks where available, while still providing access to 2G networks where this is not the case. This implies that there will be a need for multi-mode and multi-band power amplifiers as well.

1.2 Challenges for PA design

The desired functionality for these RF power amplifiers is easily described and modeled: it should accurately amplify an incoming RF signal by a fixed (or programmable) gain:

$$P_{out} = G_p * P_{in}$$

With P_{out} the output power, P_{in} the input power, and G_p the power gain.

The simplicity of this desired functionality is apparent from many modern implementations, which consist of relatively few active devices (starting at 2 transistors).

Therefore, it might seem that the design of such a simple function requires little effort and deserves little attention. As with many other RF circuits in cellular phones, the justification for all the effort that goes into their design is derived from the combination of:

- the importance of these circuits to the overall performance of the handset
- the many specifications that need to be achieved simultaneously

The power amplifier is important to the overall performance of the handset since it typically consumes the largest part of the power in a handset when active, and is therefore the most important factor in the talk time of a handset. For that reason, power efficiency is a very important specification of a PA.

This efficiency has to be achieved while still meeting the many specifications required to have the handset work well within the system:

- Linearity is becoming an important issue especially in newer systems that use advanced modulation schemes to achieve better bandwidth efficiency
- Robustness is important since the handset is part of a rather variable environment, in which power supply voltage, load impedance, temperature, transmit frequency, and output power can vary quickly and sometimes over large ranges. Since the optimization of efficiency often results in voltages and currents close to the reliability limits of the technology, significant changes in any of these parameters can result in performance degradation or even complete failure of the device. Conversely, preventing such robustness problems often results in designs with voltages and currents that cannot be optimized for efficiency.

- Stability, especially under load mismatch conditions. Such conditions can for example arise when the antenna environment changes.
- Noise, especially in the receive band of the system, since this affects the sensitivity of receivers in the system
- Spurious emissions, which can interfere with other electronic equipment or with transmissions from other handsets or from basestations in the same system
- Thermal behavior, including performance impact of temperature changes on the handset, and impact on reliability
- Multi-mode multi-band functionality, which requires adjustable properties of the power amplifier, and/or switches and adjustable circuits around a number of individual power amplifiers. This added complexity affects in turn the other specifications such as gain, linearity, output power, etc.

This paper will give an overview of these issues. In the next section, the relation between bandwidth efficiency, power amplifier linearity and power efficiency will be discussed at system and circuit level.
Section 3 describes the environment of the power amplifier, which is the basis for relating the systems considerations of the previous section (2) to the PA issues in the next section (4).

Implementing the power amplifier is discussed in section 5, using a Si/LTCC integrated GSM/EDGE power amplifier module to demonstrate the relevant issues.

2. Bandwidth Efficiency, Power Efficiency and PA Linearity

In the past more capacity could be found by moving up in frequency with improvements in device technology (recently into the 2GHz range). This trend is not likely to continue in the future because the link budget is unfavorable for such higher frequencies (fig. 1). Instead, the increased capacity is achieved by more efficient use of available bandwidth through advanced modulation schemes and access methods.

An important consequence for power amplifiers is that these efficient modulation schemes (such as QPSK, QAM, etc) are not constant envelope anymore, and therefore require more linearity from the power amplifier.

It is not likely that customers will accept significant reductions in talk time or larger and more expensive batteries and handsets, therefore the efficiency of the PA cannot be compromised too much by the new linearity requirements.

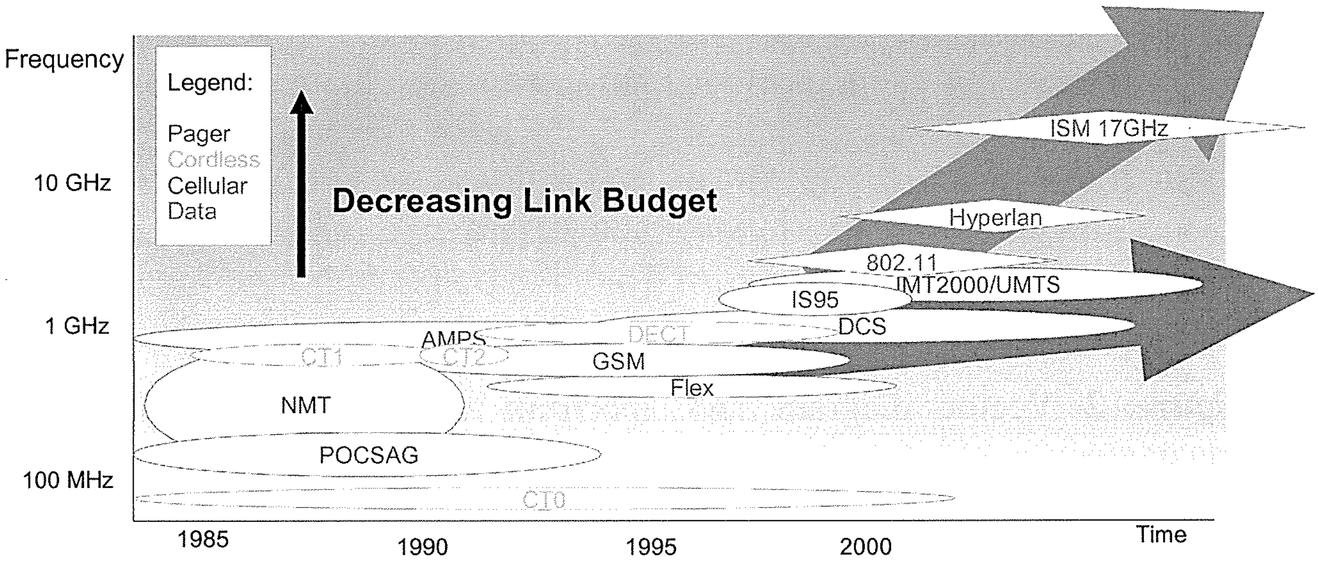

Figure 1 Frequencies used in wireless systems versus time

At the system level these amplifier non-linearities results in a so called spectral re-growth. In-band energy is transformed into energy out of band that might disturb reception in adjacent frequency channels.

Figure 2 and Figure 3 show the effect of non-linearity (in this case hard-limiting) on vector diagrams and transmission spectra (spectral re-growth) of advanced modulation schemes (in this case UMTS).

Spectral re-growth is quantified through a parameter called adjacent channel power rejection or ACPR. This parameter defines the amount of energy in the adjacent channel relative to the energy of the transmitted signal.

Non-linearities also have an impact on the wanted signal in the sense that amplitude and phase information, modulated onto the carrier, are disturbed and therefore demodulation on the receiving side can result in incorrect data at base-band. This data is often visualised as a set of amplitude normalised discrete I and Q values that represent the symbols being transmitted. Due to distortion the I and Q values of each symbol shift. The Error Vector Magnitude (EVM) is used to quantify this shift.

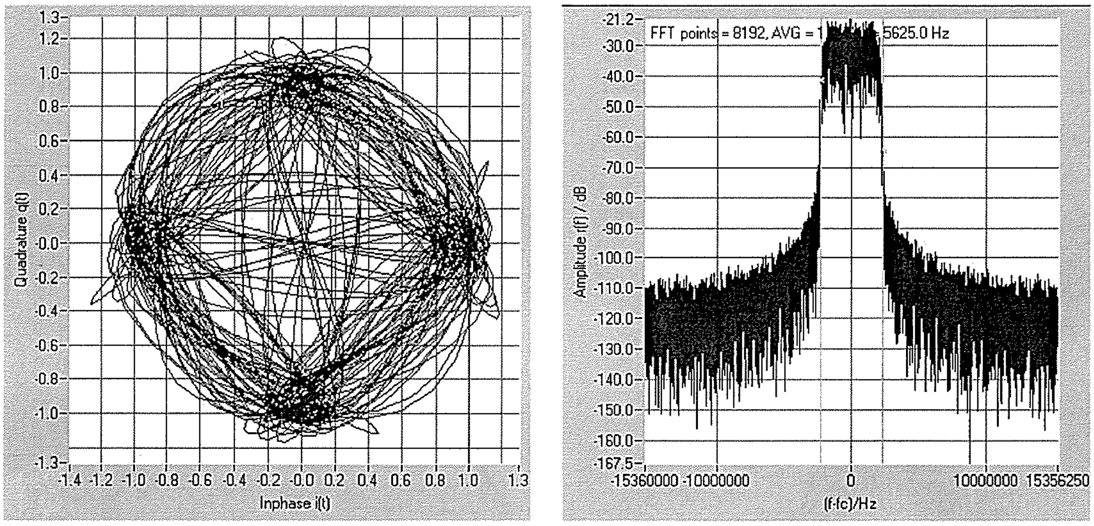

Figure 2 Vector diagram and spectrum of an ideal UMTS handset transmitter.

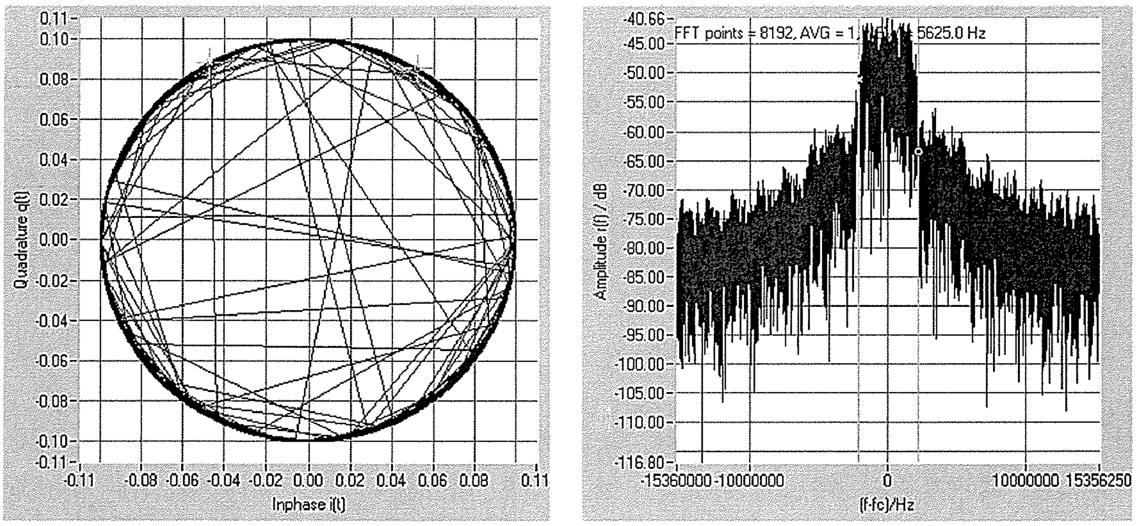

Figure 3 Vector diagram and spectrum of a hard-limited (near-constant envelope) UMTS signal

ACPR and EVM are determined by the amplifier non-linearities in combination with the signal. For the various systems both the requirements on ACPR and EVM as well as the properties of the signal (Pout, peak-to-average, power density distribution, ..) are different. This makes comparison of linearity requirements for the various standards difficult.

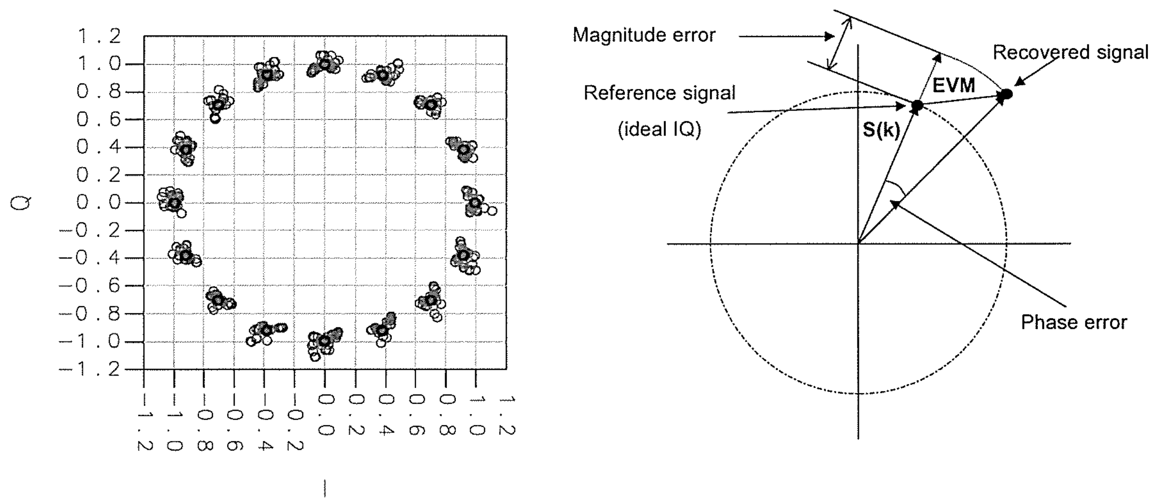

Figure 4 Due to distortion of the amplifier the symbols being transmitted, represented by their amplitude normalised I and Q values, are shifted. This is defined as Error Vector Magnitude (EVM).

On the other hand, AM-to-AM and AM-to-PM conversion are inherent properties of the circuit. For a given protocol ACPR and EVM are related to the combination of AM-to-AM and AM-to-PM. Therefore a maximum allowable AM-to-AM requirement can not be defined independent of the maximum allowable AM-to-PM and visa versa. In practice a whole set of AM-to-AM and AM-to-PM combination can be found that fulfil the ACPR and EVM requirements. A much larger set that doesn't. Consequently, PA circuit optimisation can best be done by optimising for the system parameters ACPR and EVM rather than for AM-to-AM and AM-to-PM [2].

A single stage bipolar amplifier has several causes of non-linearity. We can distinguish contributions due to voltage saturation at the collector, transistor current density variations, supply voltage variations at the transistor base and supply voltage variations at the collector.

2.1 Voltage saturation.

Power amplifiers are optimised for optimum power added efficiency and linearity. The collector load impedance at the fundamental frequency and harmonics is chosen such that at the maximum required output power the complete signal voltage headroom at the collector is used. Consequently, the transistor is driven in saturation as much as possible up to the point where saturation becomes too severe. This trade-off is limited by the emitter ballast resistors needed for thermal stability and by the transistor collector resistance and quasi saturation behaviour related to that.

2.2 Current density variations

Although power amplifiers for applications like cdmaOne, Edge, W-CDMA etc. are often referred to as linear amplifiers their behaviour is non-linear. Due to class A/B operation the current through an RF transistor, biased typically at 50mA, can increase up to an average of 500mA at maximum output power. Under influence of the carrier envelope the transistor operating point changes drastically. As a result the transistor input impedance varies with the carrier envelope [3]. The impedance match with the source is power dependent and thus varies with the envelope. This effect can be used advantageously. At power levels close to saturation the amplitude of the gain tends to drop down. When the source impedance match is optimal around this power level and less optimum at lower power levels the gain can be flattened out over a wider power level range [4].

2.3 Supply voltage variations

Non-linearity of the RF-transistor results in low frequency components in the collector and base current. These low frequency components are related to the data modulated on to the carrier. For Edge the modulation bandwidth is 100kHz approximately. Any resistance in the power supply of the collector will result in low frequency supply voltage variations at the collector. At high power levels this drives the RF-transistor further in to saturation. Therefore, proper LF-decoupling of the collector supply voltage is necessary for achieving maximum linearity.

In the base of the RF-transistor low frequency current components are present beta times smaller than in the collector. Any resistance in the voltage supply of the base (output resistance of the biasing circuit) results in low frequency supply variations at the base.
The voltage drop due to the output resistance of the biasing circuit changes the operating point of the RF-transistor which results in additional distortion. In a TDMA system, like GSM, the amplifier is switched on and off in bursts by switching the biasing circuitry. Consequently, LF-decoupling of the base can not be applied because it would disturb the amplifier turn-on and turn-off behaviour.

Together, this behavior results in a trade-off of linearity and efficiency as shown in the figure below (Figure 5):

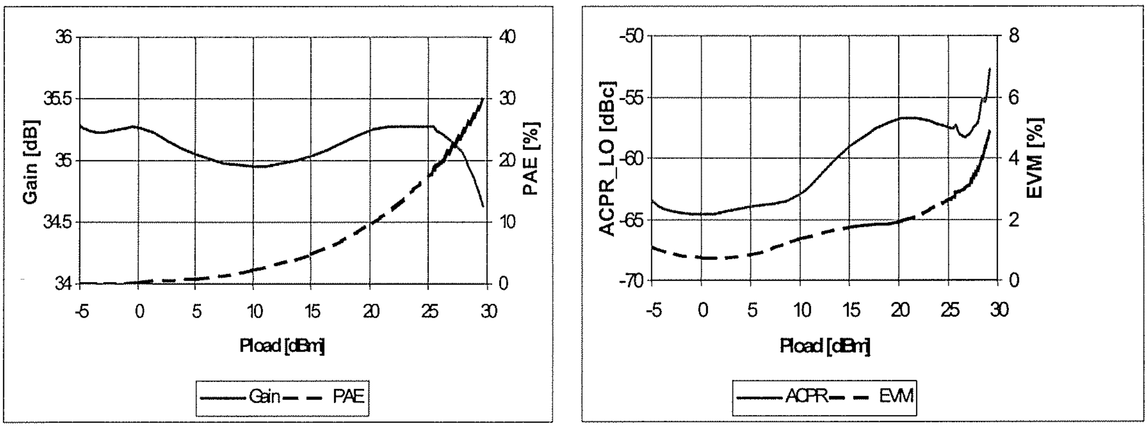

Figure 5 Gain, Efficiency (PAE), Linearity (ACPR) and EVM trade-off

3. Environment

The relation between the system requirements and the PA requirements is determined by the environment of the PA. The environment of the PA typically consists of:

- a transceiver IC at the input. This transceiver IC generates the RF transmitter signal at a low power level, often around 0dBm. A filter is often placed between the transceiver IC output and the PA input to eliminate noise from the transmitter IC outside the transmission band.

- antenna interface circuits that can include matching circuits, isolator, duplexer, switches, and diplexers to connect the PA to one or more antennas
- a control IC that sets gain, biasing, and/or output power levels through a number of control pins
- a power supply, which is often directly coming from the battery, but can also be provided through a DC/DC converter

A typical PA environment is shown in the figure below (Fig. 6):

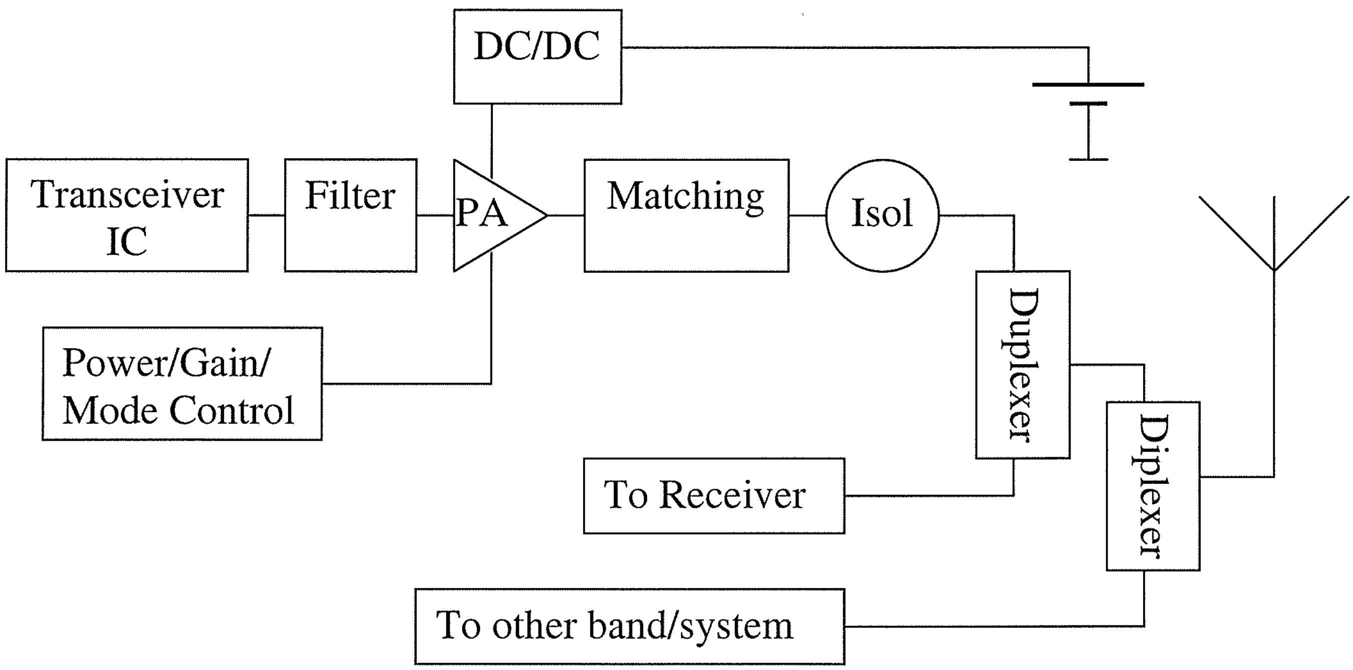

Figure 6 Environment of the PA

The duplexer is responsible for connecting the antenna to both the transmitter and receiver in such way that the energy from the transmitter is sent to the antenna only (and not the receiver), whereas the energy received by the antenna is sent to the receiver only (and not the transmitter). Depending on the access and duplex methods of the system, the duplexer can be implemented either as a traditional duplexer filter and/or through switches.
The diplexer connects transceivers for different systems and/or bands to the antenna. Again, depending on the properties of the systems, this can be

implemented through filters and/or switches.
The isolator serves to protect the power amplifier from impedance mismatches at the duplexer input. This is not always necessary, e.g. for GSM type systems this component is typically left out. By presenting the PA a fixed and nominal load impedance independent of the actual input impedance of the duplexer, the design of the PA can be further optimized since the influence of the load impedance on linearity, reliability and stability does not have to be taken into account.

To give a first impression of the type of load change that can be expected from an antenna through changes in the environment, the figure below (Figure 7) shows simulation results of a dipole with and without a conducting body at 2cm distance, not untypical for a handset antenna near the head or in free space.

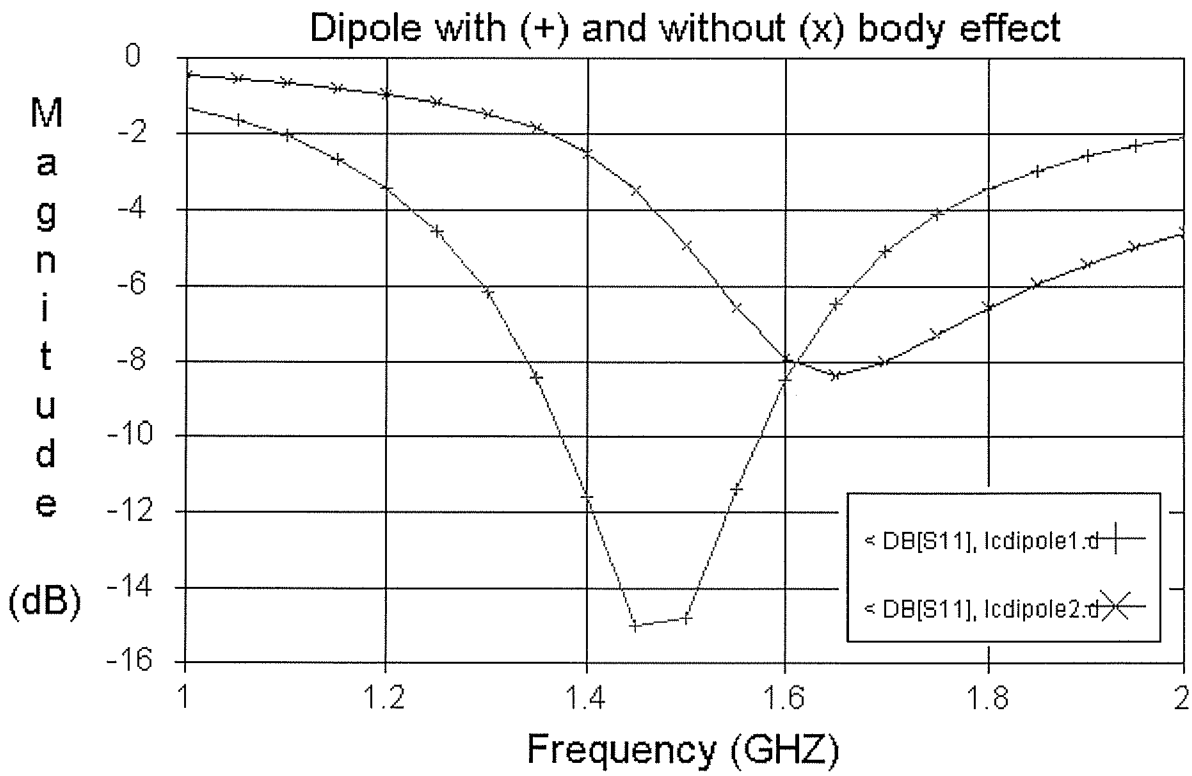

Figure 7 Antenna impedance in different environments

As shown by the simulation results, the impedance change of the antenna is quite dramatic and results in large changes of the return loss, e.g. from –10dB to –2dB around 1.37GHz. This results in a reduction of transmitted power from 90% to 37%. Considering all the effort spent on optimizing the efficiency of the PA, these numbers are very significant. Measurements on various antennas show that these numbers do occur in practice as well, and in some cases can be even worse.

Since efficiency is such an important parameter, it is very useful to find out where power is lost in the total system. The figure below shows a typical situation for a GSM PA in a multi-mode system. The numbers represent the power consumption in Watt.

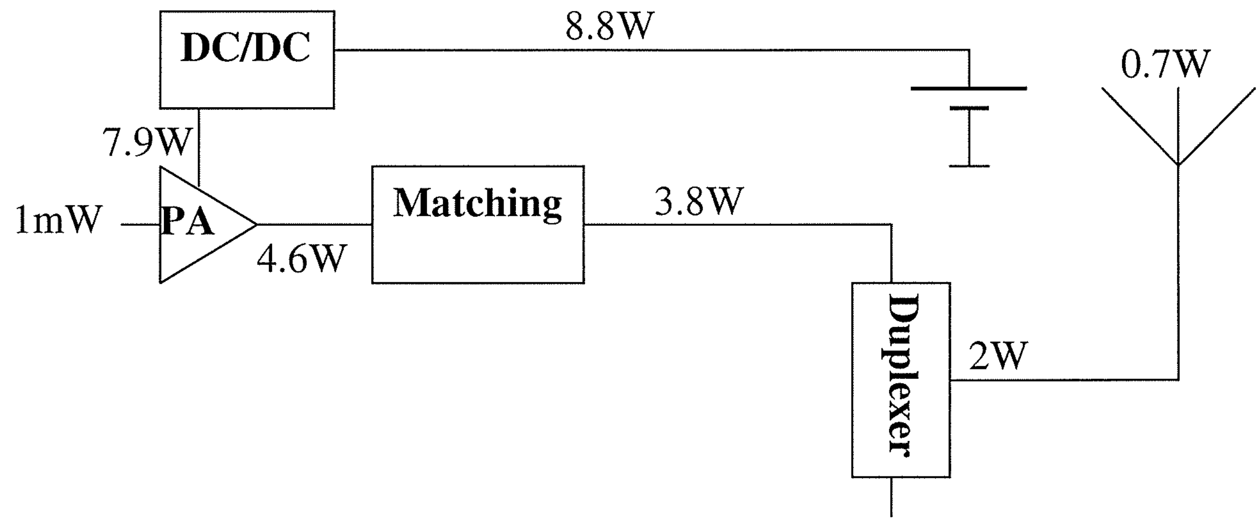

Figure 8 Power losses in the PA environment

From this figure, it becomes clear that there is a very large power loss between power drained from the battery and power ultimately delivered to the electromagnetic field: the overall efficiency in this not so unrealistic scenario is 8%, and is composed of the following major items:

- PA proper: 58% efficiency
- Antenna interface (matching, duplexer): 43% efficiency
- Antenna: 37% efficiency.

Considering that the theoretical efficiency of an ideal class A/B amplifier is

78%, it is obvious that the potential for improving overall efficiency by improving the PA proper (e.g. by going to more expensive active devices) is limited. Instead, passive devices and the antenna are more obvious candidates for overall efficiency improvement.

4. Power amplifier

After taking into account the environment of the PA, what remains is a number of issues and specifications that need to be achieved in the PA itself, through careful choices in the partitioning, implementation and technologies.

It is rather common to implement GSM power amplifiers as hybrids. This allows for usage of best combinations of active and passive technologies in order to be able to meet severe specifications on reliability, ruggedness, stability, power added efficiency, size and cost. Moreover, a hybrid power amplifier solutions is attractive because, due to the 50Ω matching networks at input and output and the on-module power supply decouplings, the amplifier function is well defined and therefore easily applicable.

Reliability (life-time) of a GSM amplifier is mainly related to the maximum temperatures that occur. Especially for the recently defined class 12 operation, with an on/off duty cycle of 50%, the solder between PCB and hybrid module, the glue to attach the die on the LTCC substrate and the Aluminium interconnect of the die might approach critical temperature values.

Moreover, the amplifier has to survive very severe conditions that might happen occasionally. For instance, the amplifier should not be damaged when the antenna is being disconnected while the battery is being charged and collector voltages up to 20V may occur. This poses rather severe requirements on the collector-base breakdown voltage.

In a GSM handset the power amplifier dissipates a significant amount of power and thus determines the standby and talk time in to a great extend. In particular the final RF-transistor geometry and the output matching network have to be designed for maximum power added efficiency [1]. The output matching network provides an optimum collector load impedance for the fundamental frequency as well as for the harmonics. It is realised by means of High-Q microstrip lines, integrated on LTCC, and high-Q SMD capacitors in order to minimise insertion losses.

For a dual mode GSM/Edge power amplifier additional requirements with respect to linearity have to be met. There is not much design freedom for optimising the linearity of a GSM/Edge amplifier when typical GSM specifications have to be met anyway. In this example, the biasing of the RF transistor in GSM mode has been made independent of the biasing in Edge mode. Optimum linearity is achieved by optimising the DC operating points of the three cascaded RF-transistors each.

5. Implementation

In this section we will discuss implementation details of the GSM/EDGE PA. This PA is relevant for a number of reasons:

- This combination of systems in a single handset is likely to become popular
- It is an optimised combination of a saturated, strongly non-linear PA for GSM mode and a linear PA for EDGE mode, with integrated mode switching
- It is typical for many of the multi-mode multi-band power amplifiers that will be needed in the transition period between 2G and 3G systems.

The Edge protocol has been adopted as an evolutionary path for enhanced data-rates and increased capacity in GSM. Edge is compatible to GSM in the sense that it operates in the same frequency bands and that it makes use of the same channel bandwidth and channel spacing. The data-rate, however, has been made a factor 3 higher by applying $3\pi/8$ offset 8-PSK (non-constant envelope) modulation and appropriate modulation filtering.
The amplifier module consists of two fully independent RF line-ups (see Figure 9). Each line-up consists of a 50 ohm input matching circuit, three

cascaded RF transistor with interstage matching circuits in between and a 50 ohm output matching circuit. The module can operate either in GSM-mode or Edge-mode by activating the GSM biasing circuits or Edge biasing circuits respectively.

In GSM mode the output power can be controlled with Vcntrl. In Edge mode, however, the output power is determined by the input power. The gain of the amplifier is constant. The biasing circuits of Edge-mode are activated by applying a stabilised voltage Vstab.

As shown in the block diagram, the module contains an output power detector for 900MHz and for 1800/1900MHz. These outputs can be used to close a power control loop for smooth up and down ramping of the power.
To study and optimise the linearity in Edge-mode the bias current of the 2nd and 3rd RF-transistor can be enhanced by applying a current Iref2 and Iref3 respectively.

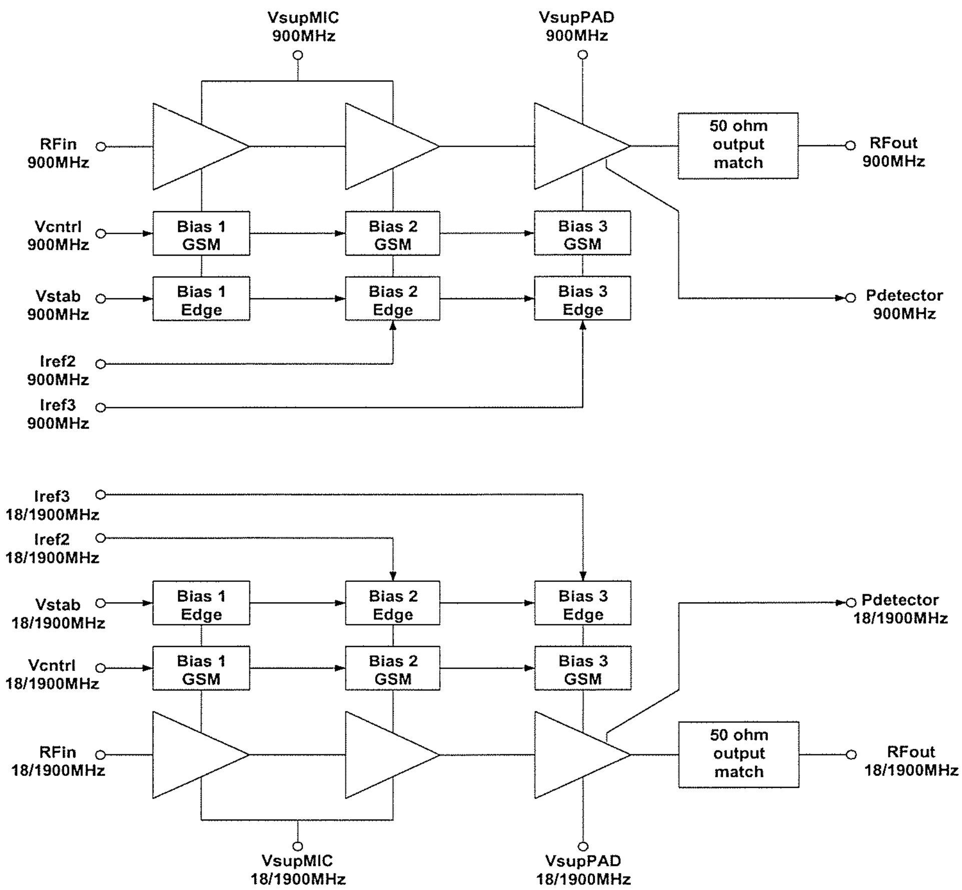

Figure 9 Block diagram of the GSM/Edge Power Amplifier module.

Figure 10 shows a photograph of the triple-band GSM/Edge power amplifier. The 900MHz line-up is visible at the left hand side and the wide-band 1800/1900MHz line-up at the right hand side. The die at the bottom side forms the driver IC for the final stage that is positioned at the top side. 0402 SMDs are used for decoupling of the power supply lines feeding the RF-transistors and biasing circuits. Input, interstage and output matching networks are build-up with discrete capacitors and microstrip line inductors on top of the ceramic substrate. In order to reduce DC and/or RF losses relatively wide traces are used for the RF-choke that feeds the final stage, as well as for the output matching microstrip lines.
In a final product the module is encapsulated with an 0.25mm thin plastic cap. The module size is 11x13.75x1.8mm.

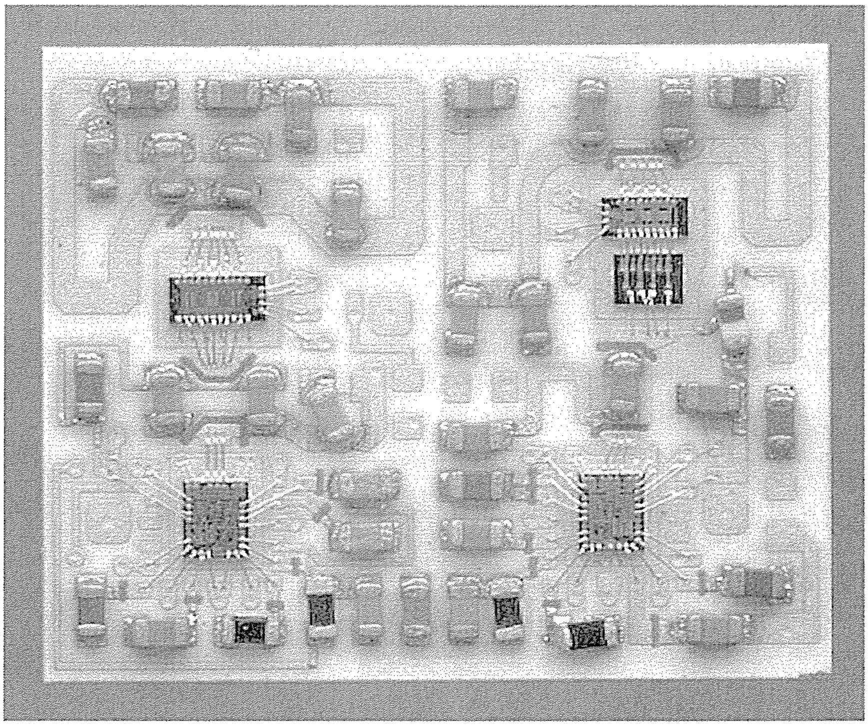

Figure 10 On the photograph of the module the 900MHz line-up is visible on the left hand side and the 1800/1900MHz line-up on the right hand side.

5.1 Output matching network

On the left top side of the module the RC-choke to feed the 900MHz final stage is visible. The choke is RF decoupled at the supply side and made resonant using a capacitor located close to the collector bondwires. The output matching network, located next right to the RF choke, consists of several sections to transform the 50ohm load, in several steps, to an optimum collector impedance of about 2 Ohm at the fundamental frequency and 0.5+10j at the second harmonic. Rejection of the second and third harmonic is obtained by series resonance of matching capacitors and their series self inductance plus via inductance which gives notches in the transfer function. In simulations a typical insertion loss of 0.8dB can be obtained. The attenuation at the second and third harmonic is typically 25dB and 35dB respectively.

5.2 Thermal design

Under nominal operating conditions the amplifier module dissipates, during the power burst, approximately 3.5W when the amplifier output power is at a maximum of approx. 3.5W. Under antenna mismatch conditions combined with high battery supply voltage the power dissipation can be even twice that value. Thermal stability is ensured by applying emitter ballast resistors.

The heat, mainly generated in the emitters of the final stage, is being spread by the 200um thick silicon die and flows through the glue toward the die attach area on top of the LTCC. Internal LTCC layers are used to partly spread out the heat horizontally. The heat flows further through the copper filled vias of the LTCC substrate towards the PCB that contains several layers of copper to further spread out the heat into the telephone set. The module is designed for a thermal resistance of less than 30 K/W in order to keep the maximum die temperature below 125°C, for a maximum mounting base temperature of 85°C and a power dissipation of 7W in the pulse with an on/off duty cycle of 25%.

5.3 Biasing circuit topology

Figure 11 shows the circuit topology for biasing the 1st RF transistor in Edge mode. The current I_{REF1} drives, via the PNP mirror T60/T61, the NPN current mirror formed by T62 and T1. Emitter degeneration resistances R60/R61 are added to increase output resistance of the PNP current mirror T60/T61 in order to reduce the supply voltage dependancy and to improve matching of this current mirror [5]. T63 improves the accuracy of the NPN current mirror factor with its current multiplication factor of beta.

The resistors R63 and R62 are added to make the topology of the left hand side of the NPN mirror equal to the topology of right hand side where R16 can be used to provides RF isolation between T1 and the biasing circuit and R1 is used to degenerate T1. Summing voltages around the loop including T62 and T1 we find

$$I_{C62}R_{62} + \varphi_T \ln(\frac{I_{C62}}{I_{S62}}) + I_{B62}R_{63} = I_{C1}R_1 + \varphi_T \ln(\frac{I_{C1}}{I_{S1}}) + I_{B1}R_{16} \qquad (1)$$

Making the assumption that $\beta_{62} = \beta_{62} = \beta_1 = \beta >> 1$ we find that

$$I_{C1} = \frac{1}{R_1 + \frac{R_{16}}{\beta}} \times \left(\left(R_{62} + \frac{R_{63}}{\beta} \right) \cdot I_{REF1} + \varphi_T \ln\left(\frac{I_{REF1}}{I_{C1}} \cdot \frac{I_{S1}}{I_{S62}} \right) \right) \qquad (2)$$

Since $I_{S1} = 18 I_{S62}$, defined by emitter area ratios, the solution to (2) is

$$I_{C1} = \frac{R_{62} + \frac{R_{63}}{\beta}}{R_1 + \frac{R_{16}}{\beta}} \cdot I_{REF1} = 18 I_{REF1}$$

which is achieved by $R_{62} = 18R_1 = 18 \times 7 = 126\,\Omega$ and $R_{63} = 18R_{16} = 18 \times 100 = 1800\,\Omega$

As a result the last term in (2) goes to zero which makes the biasing of T1 almost temperature independent.

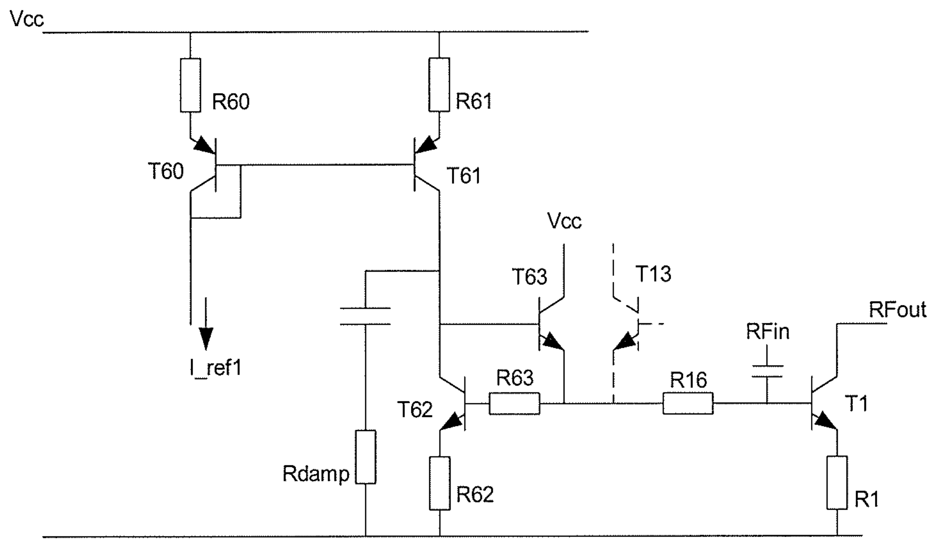

Figure 11 The RF-transistors are accurately biased with a current mirror circuit topology.

To guarantee stability of the circuit a resistor Rdamp is added to give at RF frequencies resistive loading at the high ohmic point. The dotted transistor T13 is part of the GSM biasing circuit and is not active in Edge mode.

Conclusions

Power amplifiers are very relevant components of a handset transmitter, since they consume a large part of the total power dissipation. Moreover, the linearity requirements in newer systems are difficult to combine with high efficiency. The total performance depends for a larger part on the passive components and antenna than on the active part. The difficulty in designing a good PA is in achieving good efficiency while meeting many other specifications (stability, reliability, linearity, gain, power, etc.) simultaneously.

The GSM/Edge power amplifier module used as an example throughout this paper illustrates that multi-mode multi-band power amplifiers can be realised well in Si-bipolar technology combined with multi-layer LTCC substrate.

Acknowledgements
This paper is based on insights built up in several teams throughout Philips, including the Philips Semiconductor PA development teams in Sagamihara (Japan), Mansfield (U.S.A.) and Nijmegen (The Netherlands), as well as the Philips Research Integrated Transceiver group in Eindhoven (The Netherlands).
The GSM/Edge power amplifier module could only be realised with the help from enthusiastic team members. Being aware that their contributions to the success of this project were very essential I would like to thanks Dima Prikhodko for his work on IC circuit development, Skule Pramm and Gerd Kahmen for designing the substrate and optimising the module and Christophe Chanlo for simulating ACPR and EVM. Last but not least I would like to thanks Reza Mahmoudi for the enlightened discussions we had and for developing dedicated ACPR and EVM simulation tools.

References:
[1] F. van Rijs, R. Dekker, H.A. Visser, H.G.A. Huizing, D. Hartskeerl, P.H.C. Magnee, R.Dondero. "Influence of output impedance on power added efficiency of Si-bipolar power transistor" International microwave symposium digest, Volume 3, June 11-16, 2000

[2] Private communication with Reza Mahmoudi

[3] Keng Leong Fong and Robert G. Meyer, "High-Frequency Nonlinearity Analysis of Common-Emitter and Differential-Pair Transconductance Stages", IEEE journal of solid-stade circuits vol 33, no 4, April 1998

[4] R. Mahmoudi, "Multi-Disciplinary design method for 2.5 generation of mobile communication systems" to be published in September 2001, Twente University Press.

[5] Paul R.Gray and Robert G. Meyer, "Analysis and Design of Analog Integrated Circuits", second editions, John Wiley & Sons 1984