A NEW LINEARIZATION METHOD FOR CANCELLATION OF THIRD ORDER DISTORTION

by

KONRAD MIEHLE

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in the
Department of Electrical and Computer Engineering

Charlotte

2003

Approved by:

_____
Dr. Thomas P. Weldon

_____
Dr. Farid Tranjan

_____
Dr. H. Cherukuri

ABSTRACT

KONRAD MIEHLE. A New Linearization Method for Cancellation of Third Order Distortion (Under direction of DR. THOMAS P. WELDON)

This thesis presents a new linearization method that offers the potential for reducing size, power, and cost of wireless devices such as cellular phones. In the new method, a less linear amplifier is combined with a more linear amplifier in such a manner that the combined performance greatly exceeds the performance of the two original amplifiers. This unexpected result is achieved in a simple architecture that lends itself to straightforward integrated circuit implementation without the need for bulky or expensive external components. To gain a better understanding of prior linearization techniques, an overview of function, advantages and disadvantages of prior linearization techniques is presented. The simplicity of the new linearization method is shown with a simple mathematical relationship for cancellation of third order distortion in terms of gain and third order output intercept points of the two amplifiers. In order to prove the simple mathematical relationship, a cross-coupled differential pair was implemented as a fully integrated circuit in a 0.5 μm complementary metal oxide semiconductor (CMOS) process. Measured results show very good improvement for low input power levels. The prototype was tested for different frequencies and shows an improvement of 23 dB of third order distortion for a frequency of 110 MHz. Harmonic balance simulations with Agilent Advanced Design System (ADS) were performed and show good agreement with the measured results. The new linearization method also works for devices other than amplifiers such as mixers.

ACKNOWLEDGEMENTS

I would like to thank the people who made it possible for me to study and work at the University of North Carolina at Charlotte. I would like to thank Dr. Weldon for his guidance. He gave me the opportunity to work independently on always challenging projects. I am thankful to Dr. Tranjan and Dr. Cherukuri for supporting me and being members of my committee.

Konrad Miehle

July 2003

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACPR	Adjacent Channel Power Ratio
ADS	Advanced Design System
AM/AM	Amplitude Modulation to Amplitude Modulation conversion
AM/PM	Amplitude Modulation to Phase Modulation conversion
CALLUM	Combined Analog Locked Loop Universal Modulator
CDMA	Code Division Multiple Access
CMOS	Complementary Metal Oxide Semiconductor
DSP	Digital Signal Processing
EER	Envelope Elimination and Restoration
FET	Field Effect Transistor
IF	Intermediate Frequency
IIP2	Second order Input Intercept Point
IIP3	Third order Input Intercept Point
IMD3	Third order Intermodulation Product
IP	Intercept Point
IP2	Second order Intercept Point
IP3	Third order Intercept Point
IP1dB	Input 1dB Compression Point
LINC	Linear amplification using Nonlinear Components
LNA	Low Noise Amplifier
LO	Local Oscillator

MOS	Metal Oxide Semiconductor
OP1dB	Output 1dB Compression Point
OIP2	Second order Output Intercept Point
OIP3	Third order Output Intercept Point
P1dB	1dB Compression Point
PA	Power Amplifier
PLL	Phase Lock Loop
QAM	Quadrature Amplitude Modulation
QPSK	Quadrature Phase Shift Keying
RF	Radio Frequency
SSB	Single-Sideband
VCO	Voltage Controlled Oscillator

CHAPTER 1: INTRODUCTION

Despite considerable advances in technology, nonlinearity of devices continues to limit the performance of wireless systems and devices. In the design of radio receivers, nonlinearity restricts the ability of a radio to receive weak signals in the presence of nearby stronger signals [1]. In radio transmitters, nonlinearity can cause the transmitted signal to spill over into adjacent frequency channels, interfering with other users [1]. Therefore, the present thesis presents new approaches to reduce nonlinearity in devices and systems.

In many applications, a specific nonlinearity known as third order nonlinearity causes particular problems. When a signal encounters such a third order nonlinearity, the resulting signal distortion is referred to as third order intermodulation distortion. It is this third order intermodulation distortion that often causes greatest difficulty in radio transmitters and radio receivers. To overcome this problem, transmitters and receivers are often designed with higher levels of power consumption. However, such increased power consumption is obviously undesirable in portable devices such as cellular phones. In addition, larger and more expensive devices usually accompany the higher power level. The approach presented in this thesis also seeks to mitigate such limitations.

In the case of radio receivers, third order nonlinearity causes weak signals to be blocked in the presence of strong signals. In particular, two strong signals at frequencies adjacent to a desired weak signal can interfere with the reception of the weak signal

because of the third order nonlinearity. For example, a pair of strong signals at 100 and 101 MHz can interfere with the reception of a weak signal at 102 MHz, despite the fact that all three signals are on different frequency channels. This situation is common in cellular phone systems where nearby cellular phones may present strong signals while a user is trying to receive a weak signal from a distant radio or base station. The linearization approaches in this thesis also address these problems.

In the case of radio transmitters, third order nonlinearity causes transmitted signals to splatter into adjacent channels. Such splattering causes interference with adjacent radio channels that may be allocated to other users. In particular, a strong transmitter signal can be distorted such that it spreads into an adjacent channel. For example, a television broadcast on channel nine may inadvertently spill undesired transmitter power into channel eight or ten. Again, the present thesis addresses such radio transmitter problems.

For both transmitters and receivers, common approaches to improving linearity often require bulky components such as delay lines. Although such solutions may be acceptable for large installations such as cellular base stations and towers, these large components have obvious disadvantages for small portable applications such as cellular phones. Therefore, another intention of the thesis is to develop approaches and methods that can be implemented on integrated circuits without the need for bulky external components. Such integrated circuit implementations can be particularly important in reducing cost of high volume products such as cellular phones.

In addition to the technical problems addressed by this thesis, a simple analytic approach to linearization problems is presented. The simple approach is based on radio

frequency concepts of third order intercept point and results in simple algebraic equations to cancel nonlinearities. Earlier approaches commonly use cumbersome polynomial expansions and typically provide little insight for the purposes of designing new approaches to linearization. Importantly, the simple analysis methods give greater insight for the development of new linearization architectures. Furthermore, the analysis is not dependent on underlying assumptions such as square-law metal oxide semiconductor field effect transistor (MOSFET) behavior.

Finally, an unexpected result of the present thesis is that a "bad" device can be added to a "good" device to form a composite device that is much better than either of the two original devices. In more technical terms, a very nonlinear device can be added to a slightly nonlinear device to form a composite device that is much more linear than either of the two original devices. Although earlier researchers used a similar cross-coupled differential pair method, the present work provides the foundation for new four-stage approaches with advantages over the cross-coupled differential pair [21], [37].

1.1 Thesis Statement

This thesis focuses on new linearization methods for cancellation of third order intermodulation distortion. A simple analytic approach is presented which promises to be a basis for various new linearization techniques. The analytic approach has a simple algebraic relationship, leading to a wide variety of new ideas for integrated circuit implementations. The new method is applicable not only to radio transmitters, but also to radio receivers. Furthermore, it can be fully integrated and does not require additional external devices, such as delay lines.

1.2 Organization

Chapter 2 describes prior linearization methods including feedback, feedforward, and predistortion methods. Chapter 3 presents theory and descriptions of nonlinear circuits as well as important concepts such as third order intercept and compression point. The new method of cancellation of third order nonlinearity is described in Chapter 4. Chapter 5 presents the cross-coupled differential pair implemented as integrated circuit and discusses measured results. Measured and simulated results are compared in Chapter 6. Finally, Chapter 7 summarizes this work and presents preliminary results for future research. Part of this future research is a fully integrable four-stage implementation, which promises to cancel third order distortion up to the 1dB compression point.

CHAPTER 2: PRIOR LINEARIZATION METHODS

In this chapter, prior research on linearization methods is first reviewed before describing theoretical background material in Chapter 3. These prior approaches provide the context for development of the new methods of linearization in Chapter 4. In the following, prior linearization methods are roughly split into three categories: Feedback, Feedforward, and Predistortion [3].

2.1 Feedback Linearization

The first general category of linearization is Feedback linearization. Feedback linearization linearizes transmitters or single PA's by forcing the output to follow the input. The Feedback technique can be either directly applied to the RF signal or indirectly to the modulation, i.e. envelope, phase or I (in-phase) and Q (quadrature) components [1]. In the following section, RF feedback using direct feedback is first discussed. Then, Envelope feedback is described where the envelopes of input and output signals are extracted and compared. Finally, Polar and Cartesian loop feedback are discussed, where the linearization scheme includes the complete transmitter.

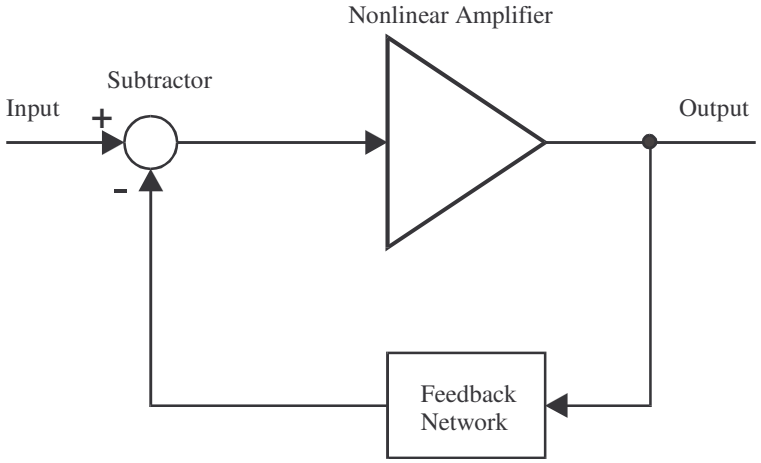

Figure 2.1 RF feedback linearization method: output of nonlinear amplifier is fed back and subtracted from the input signal. Feedback network is typically a constant attenuation network.

2.1.1 RF Feedback

In radio frequency (RF) feedback the output signal is fed back without detection or down-conversion. RF feedback is illustrated in Fig 2.1. The RF signal is input to a subtractor on the left side in Fig. 2.1. The output signal of the amplifier, on top of Fig. 2.1, is fed back to the subtractor and subtracted from the RF input signal. The feedback network, at the bottom of Fig. 2.1 can either be passive or active. An amplifier can be used for an active feedback network or resistors or transformers can be deployed as passive feedback networks. The feedback network can reduce distortion appearing at the output of the nonlinear amplifier in Fig. 2.1.

Voltage-controlled current feedback and current-controlled voltage feedback are commonly used for this method because they are simple and their distortion performance is predictable [1]. However, due to the time delays in the feedback network, loop stability is a problem in this design. As a result, RF feedback is limited to narrowband systems. A

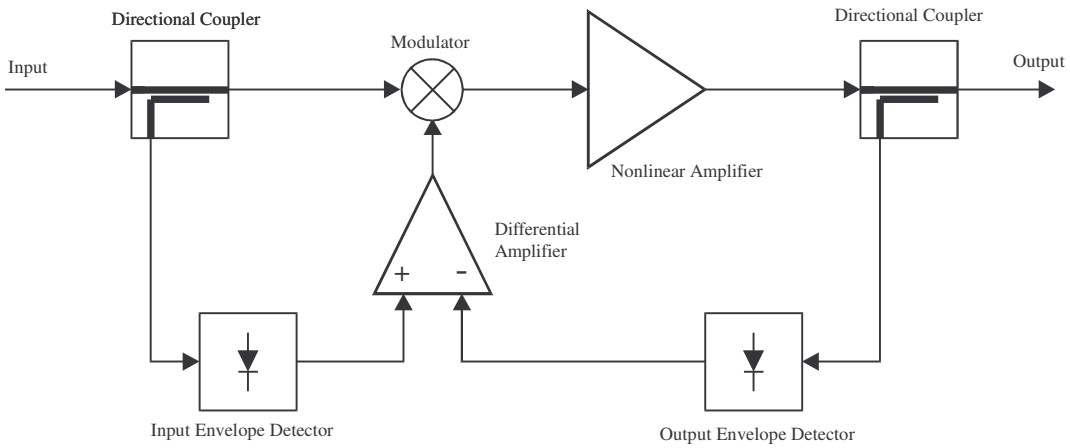

Figure 2.2 Envelope feedback linearization: on the left side an input coupler splits the input signal. The envelope detector at the left detects the changes in envelope of the input signal. On the right side, the output of the amplifier is sampled by a coupler. The sample is fed into an envelope detector and the changes in output envelope are detected. A differential amplifier compares the two envelope signals and controls a modulator. The modulator adjusts the input signal to the amplifier.

loss of gain is also an issue in this technique, particularly in transmitters, where high output power is desired [1], [4].

2.1.2 Envelope Feedback

A second feedback linearization method is envelope feedback. Fig. 2.2 shows an envelope feedback scheme. In Fig. 2.2, a portion of input and output signals is sampled by a coupler at the input and output and the envelopes of the two sampled signals are detected by means of an envelope detector at the bottom of Fig. 2.2. Both envelope signals are subtracted using a differential amplifier. The resulting error-signal controls a modulator, which modifies the envelope of the RF input signal. The output signal of the modulator is amplified by the nonlinear power amplifier.

In envelope feedback, both modulator and PA are included in the linearization process. Furthermore, envelope feedback can be applied to either a transmitter or a single

PA. However, a disadvantage of the envelope feedback is that it only accounts for distortion in signal-amplitude and not in signal-phase [1], [4].

2.1.3 Polar Loop Feedback

Polar Loop is an improved version of envelope feedback. The Polar Loop scheme compensates amplitude and phase distortion. In most cases, envelope and phase comparison takes place at the intermediate frequency (IF). Typically, the polar loop technique is deployed to a complete transmitter rather than a single PA [1], [3], [4], [5], [6].

A Polar Loop linearization technique is illustrated in Fig. 2.3. The operation of a polar loop is similar to envelope feedback. A sample of the output signal is down-converted to a convenient IF by the local oscillator (LO) at the right bottom corner of Fig. 2.3. The IF signal is then separated into phase and amplitude by a limiter and demodulator, respectively. This procedure is illustrated at the bottom center in Fig. 3.2. On the left side in Fig. 2.3, a single sideband (SSB) modulated signal, as input signal, is split and separated into phase and amplitude by a limiter and demodulator, respectively, similar to the sampled output signal. In the center of Fig. 2.3, both output and input amplitudes are compared with an error amplifier and the resulting error signal controls a modulation amplifier. The phase signals of input and output are multiplied utilizing a mixer shown at the bottom of Fig. 3.2. The resulting signal controls a voltage-controlled oscillator (VCO) after passing a loop filter and being amplified. The new-formed phase and amplitude error signals are combined with a modulating amplifier at the top of Fig. 3.2. Finally, the combined signals are amplified with a power amplifier. In Fig. 2.3, the loop, which determines the phase error-signal, is essentially a phase lock loop (PLL) with

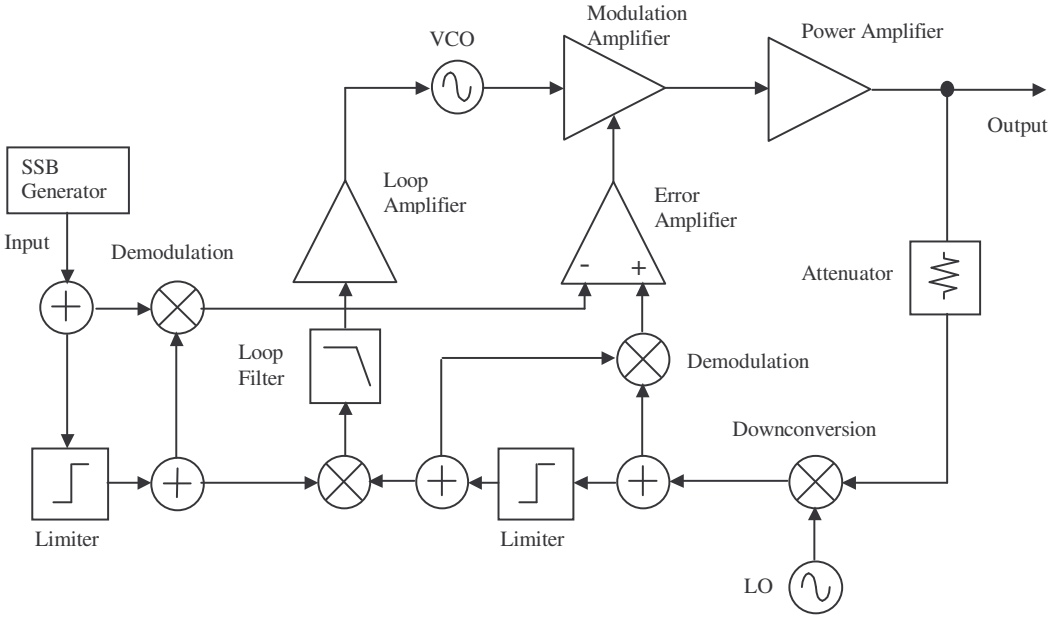

Figure 2.3 Polar loop feedback linearization: on the left side, the input is a single sideband modulated (SSB) signal which is split and input to a limiter and demodulator. On the right side, the output signal is down-converted with a local oscillator (LO), split, and also input to a limiter and demodulator. The amplitude information is output of the demodulators and is fed to an error amplifier. The error-signal controls a modulation amplifier. The output signals of the limiters are split and fed back to the demodulator. The second outputs of the splitters are compared using a mixer. The output signal of the mixer is filtered and amplified with a loop amplifier and controls a voltage controlled oscillator (VCO). The amplitude modulated output signal of the VCO is input to the power amplifier [1].

a mixer as phase detector, a VCO and an amplifier in the loop [5]. A disadvantage of the PLL is the locking problem at low amplitudes [5]. The PLL is also an example of the complexity of the polar loop technique. Another disadvantage of the polar loop linearization is that the extracted input and output phase bandwidth must be at least ten times the RF bandwidth at the output [3].

2.1.4 Cartesian Loop Feedback

Another feedback linearization technique is the Cartesian Loop feedback which is depicted in Fig. 2.4. In this technique, the input and output baseband signals shown on the left in Fig. 2.4 are processed in Cartesian rather than polar form. The input baseband signals are available as quadrature components, I and Q, on the left side in Fig. 2.4. Both signals are fed into differential amplifiers and are up-converted to RF using a quadrature amplitude modulator (QAM). The signals at the output of the QAM up-converters are

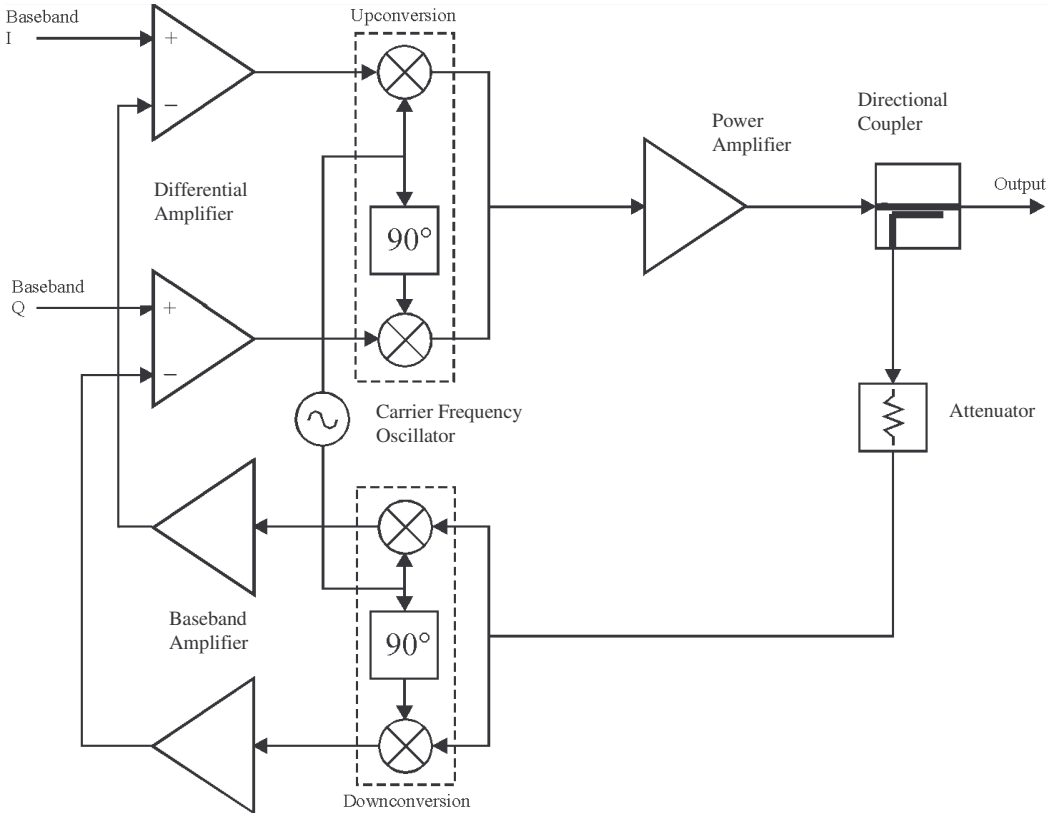

Figure 2.4 Cartesian loop feedback linearization: on the left side the input signals are available as I and Q components and are input to differential amplifiers. On the right side the output signal is sampled. The sampled signal is, after attenuation, fed to a quadrature amplitude modulator (QAM) where the sample is down-converted into I and Q components. Baseband amplifiers amplify the I and Q components and close the loop to the differential amplifiers. The output signals of the differential amplifiers are up-converted with a QAM and amplified with a power amplifier. The QAM modulator consists of two mixer and 90 degree phase shift network to create in-phase and quadrature components [1].

combined forming a complex RF signal, which is amplified by a nonlinear PA. To close the loop to the differential amplifiers, the output signal on the right side in Fig. 2.4 is sampled, down-converted and resolved into components I and Q [1], [3], [4], [5]. The QAM modulators typically consist of two mixers and a 90° phase shift network.

The Cartesian Loop feedback forms a complete linear transmitter. Nonlinearities of all blocks, such as the QAM up-converter in the loop, are removed. However, as in all feedback systems, the degree of linearity improvement depends on the delay around the loop [1]. Hence, its performance depends on the bandwidth of the feedback system. Higher frequency components experience more time delay in the feedback system than low frequency components. Thus, the feedback system can become unstable at signals with high frequency components and cause unwanted oscillations. To prevent such unstable behavior, feedback systems are limited in bandwidth.

Faulkner [7] reports a Cartesian feedback, which employs a phase correction circuit to stop instability and limit the level of the out-of-band noise floor. This scheme uses the supply current to the final stage amplifier as a measure of output power. The author also shows that the phase correction settings correspond to the output power condition of the amplifier. However, the correction performance is limited by hysteresis in the supply current feedback signal, which adjusts the phase corrector in the feedback path. Furthermore, Lee et al [8] utilizes a regulator for accurate oscillator phase adjustment in Cartesian feedback. However, both schemes in [7] and [8] need additional circuitry which increases overhead, complexity and cost.

2.2 Feedforward

The second major linearization method is Feedforward. In contrast to feedback

systems, feedforward linearization systems correction is not dependent on past events. Feedforward linearization is based on present events in the system and linearization is independent of amplifier delays. Additionally, basic feedforward systems are unconditionally stable and ideally do not reduce amplifier gain [1], [9], [10].

Fig. 2.5 presents a schematic of a basic feedforward system. The linearized PA with feedforward correction consists of a main and error amplifier, directional couplers, and delay lines. The incoming signal is split into two paths with one path going to the main amplifier, and the other path going to a delay element, shown in Fig. 2.5 on the right side. The signal at the output of the nonlinear amplifier contains the desired and distortion components. The delayed input signal is subtracted from a sample of the output signal of the nonlinear amplifier. The sample of the output signal of the nonlinear amplifier is taken by means of a coupler at the top of Fig. 2.5. Ideally, the carrier is canceled and the signal at the output of the subtractor contains only the distortion component. The error signal is amplified by an amplifier depicted at the bottom of Fig. 2.5 and combined with a delayed version of the output signal of the main amplifier. This second combination ideally cancels the distortion components of the main amplifier and leaves the desired signal unchanged. The first loop cancels the carrier and the second loop reduces the distortion component.

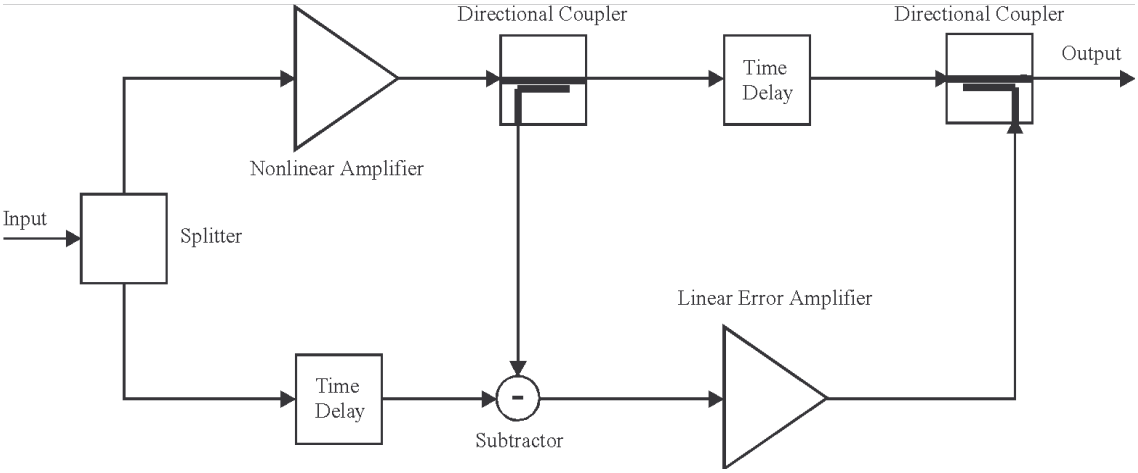

Figure 2.5 Basic feedforward linearization: on the left side the input signal is split by a splitter which is usually a passive network. The split signals are input to the main amplifier and a subtractor after being delayed. The time delay compensates the time delay of the main amplifier. A sample of the output of the amplifier is extracted with a coupler and fed to the second input of the subtractor. The output of the subtractor is input to an error amplifier on the right side. A directional coupler combines the output of the error amplifier and the delayed output of the main amplifier. The time delay on the right side is compensating the time delay of the error amplifier. The first loop eliminates the carrier frequency; the second loop cancels distortion products in the main path [9].

Feedforward can be used over a wide bandwidth of about 10-100 MHz [3]. However, inaccurate matching of devices in amplitude and phase can impair the performance of the feedback system. Since this system is feedforward in nature, alterations over aging and temperature degrade the correction of linearity. A way to reduce these effects is to use multiple feedforward loops [1]. A single feedforward loop can be considered as the main amplifier. However, by using this configuration, the complexity of a feedforward system increases fast. In addition, external devices are necessary such as delay lines.

Woo et al [11] propose a new adaptation scheme for the feedforward linearization technique. The proposed scheme uses imperfect cancellation at the error amplifier to

account for distortion created in the error amplifier. However, the proposed new adaptation scheme introduces feedback into the feedforward linearization technique, which in turn can cause stability and bandwidth problems.

2.3 Predistortion

The third major linearization technique is predistortion. Predistortion is based on the idea of inserting a nonlinear element prior to a RF PA such that the combined transfer characteristic of the two devices is linear. Fig. 2.6 shows the underlying principle. The transfer-characteristic of the predistorter, shown on the right side in Fig. 2.6, must be complementary to the transfer-characteristic of the PA, depicted in the center in Fig. 2.6. The combined transfer characteristic of the linearized amplifier is illustrated on the right in Fig. 2.6. Predistorter and amplifier transfer characteristic ideally compensate each other creating a linear input-output characteristic. Predistortion can be accomplished at either RF, IF or baseband and has the ability to linearize the entire bandwidth of an amplifier or system [3], [4]. There is also the possibility to cascade the distortion element at the output of the PA. This case is called postdistortion [1], [5]. RF/IF predistortion, which is described first, does not use feedback. Next, adaptive baseband predistortion using feedback is discussed.

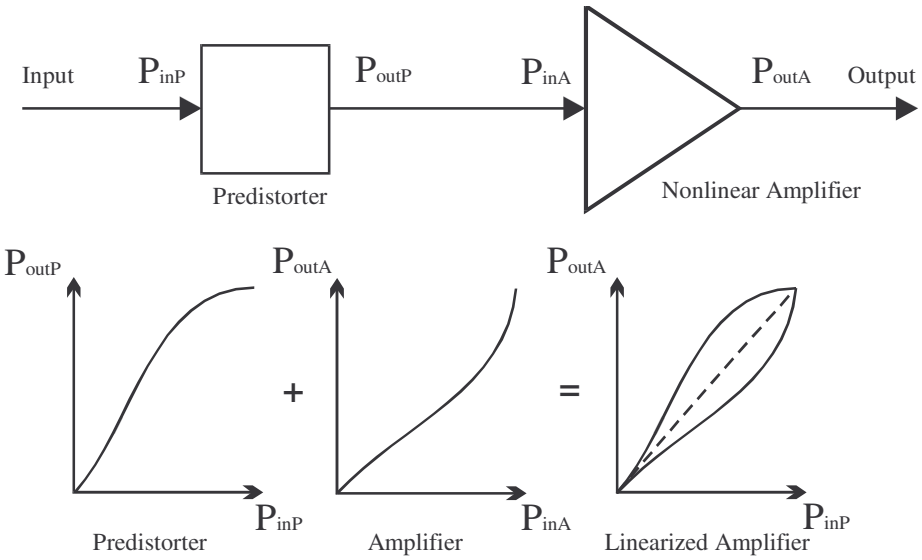

Figure 2.6 Predistortion linearization: on the top, the amplifier is cascaded with a predistorter. The predistorter can be a passive diode network or a digital signal processor. On the bottom, the transfer characteristics of the predistorter, (on the left) is added to the transfer characteristic of an amplifier (in the center). The right side shows the combined transfer characteristic, which is ideally linear.

2.3.1 RF/IF Predistortion

RF and IF predistortion are similar in operation and do not have a feedback path. Therefore, the nonlinearity to be canceled must be known in advance. In RF predistortion, the predistorter operates at the final carrier frequency. IF predistorters work at an intermediate frequency, thereby making it possible to use the predistortion network for different carrier frequencies by adjusting the local oscillator (LO) frequency [9]. The problem in this scheme is to design and fabricate a predistortion circuit, which closely resembles the required function of Fig. 2.6. The transfer characteristic of a predistortion network must be tuned to every single PA, even PA's with identical design [1].

Typically, the predistorter attempts to distort only the third-order characteristic. This type of nonlinear network is named cubic predistorter. These circuits make use of the nonlinear characteristics of devices such as diodes and transistors [12]. However, main disadvantages of RF/IF predistortion are relatively modest linearity improvement, and an inability to deal with many orders of distortion [1].

2.3.2 Adaptive Baseband Predistortion

A second type of predistortion is called adaptive baseband predistortion (also called digital predistortion). Digital predistortion is basically a Cartesian feedback with digital signal processing (DSP) added. This technique linearizes complete transmitters. Fig. 2.7 displays an adaptive baseband predistortion scheme. Predistortion occurs at the baseband level and manipulates usually I and Q components. On the right side of Fig. 2.7, the output signal is sampled, down-converted by a QAM down-converter (into components I and Q), and input to a digital signal processing (DSP) unit. Similarly, input I and Q components are directly fed into the DSP block. In order to create predistortion, weighting coefficients are stored in lookup tables in the DSP section. These coefficients can be updated by new coefficients derived from the fed back in-phase and quadrature components. The output of the DSP is the predistorted signal. This predistorted signal is then up-converted by a QAM and amplified with a PA illustrated at the top of Fig. 2.7.

The primary disadvantages of digital predistortion are its relative complexity and bandwidth limitations tied to the accuracy and computational rate of the specific DSP [13]. Furthermore, power consumption is increased due to the digital signal processor. In addition, digital predistortion exhibits storage and processing overhead for the lookup tables [1], [3].

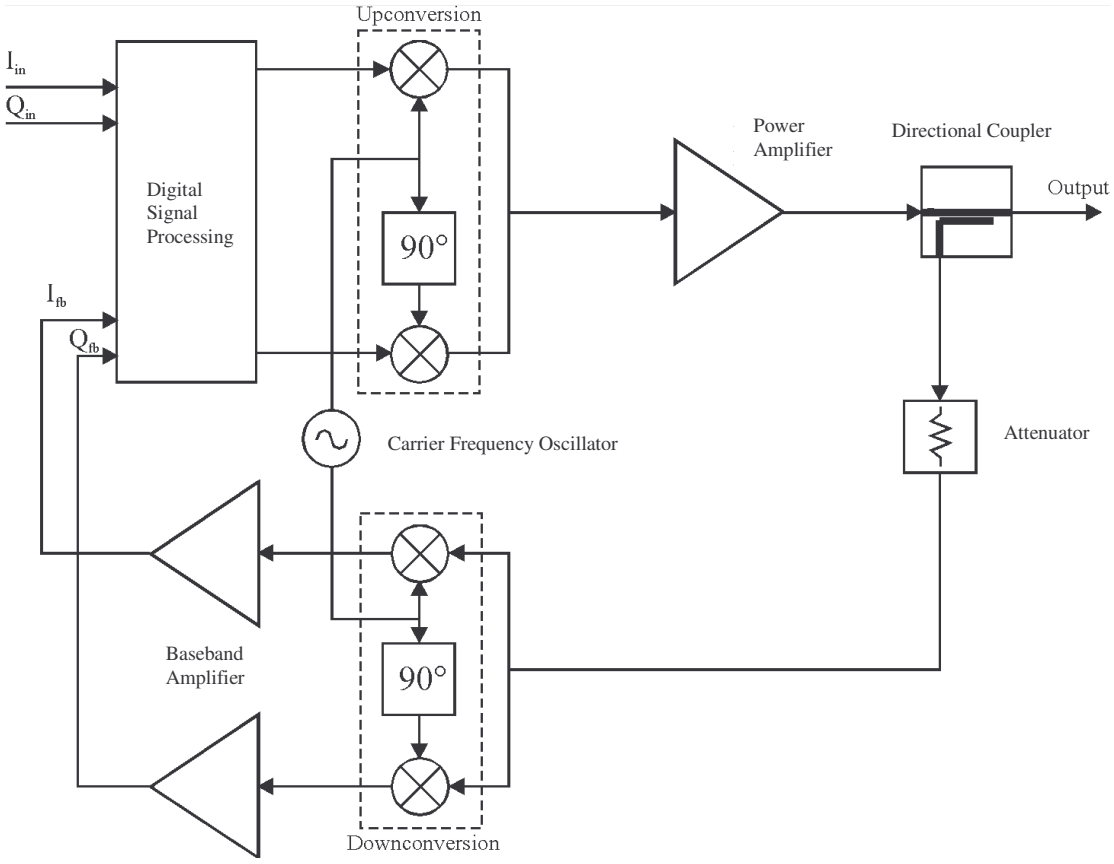

Figure 2.7 Adaptive baseband predistortion linearization: on the left side the input signals, as I and Q components are input to a digital signal processor. On the right side a sample of the output signal is taken by a coupler and after attenuation down-converted into I and Q components. The I and Q components are amplified with baseband amplifiers. The output of the amplifiers close the loop to the digital processor. The predistorted output of the digital processor is up-converted and amplified with a power amplifier. Down and upconversion is performed with a QAM converter. The QAM consists of two mixers and a 90 degree phase shift network [5].

2.4 Envelope Elimination and Restoration (EER)

This section discusses envelope elimination and restoration (EER). EER is a technique which tries to increase linearity and power efficiency simultaneously. The basic method of EER is to disassemble the RF input signal into phase and envelope components and combine both after amplification. EER is based upon the principle that

any narrow-band signal can be produced by simultaneous amplitude (envelope) and phase modulation [3]. The schematic in Fig. 2.8 shows an EER technique with limiter and envelope detector to extract the phase and envelope information, respectively. The limiter, shown at the bottom left of Fig. 2.8, eliminates the envelope and thus makes it possible for a high-efficient nonlinear PA to amplify the constant-envelope signal. Finally, the envelope amplifier (class-S modulator), depicted on the top of Fig. 2.8, modulates the final RF power amplifier and creates an amplified replica of the input signal at the output.

EER can be employed to a complete transmitter or a single PA. If EER is used to design a transmitter, a DSP is typically utilized to generate the envelope and phase information. EER reaches good linearity with high efficiency.

However, difficulties in the design lie in the limited bandwidth of the class-S modulator and correct alignment of the envelope and phase signals. Furthermore, large envelope variations can drive the RF power amplifier into cutoff causing significant distortion [1].

Methods of implementing a feedback scheme as a part of the EER technique are introduced in [14] and [15]. In the proposed feedback procedure, envelopes of input and output signals are detected and compared with a differential amplifier. The output of the differential amplifier then modulates the final stage amplifier. However, the resulting feedback poses a problem concerning bandwidth limitation and loop stability. Furthermore, complexity of the circuit topology is increased.

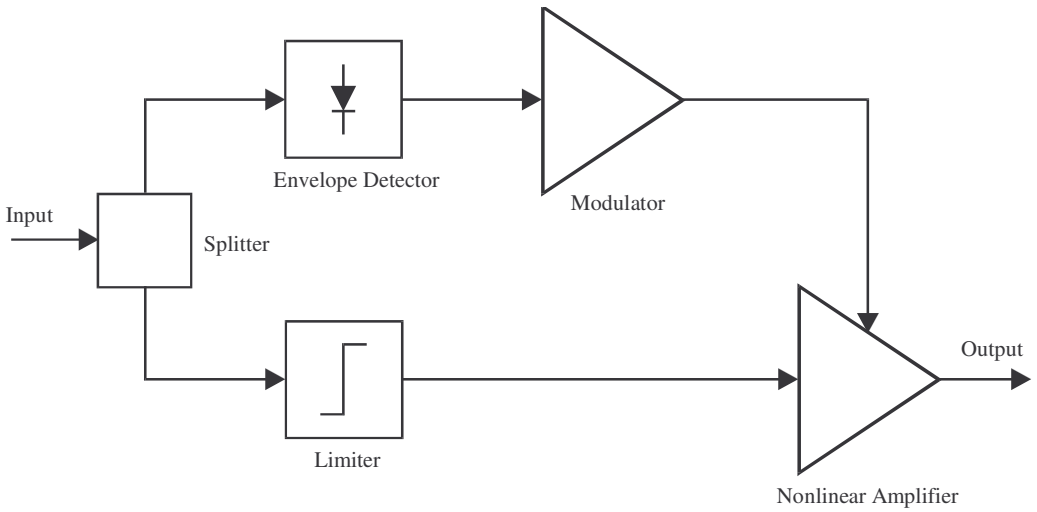

Figure 2.8 Envelope elimination and restoration: on the left, the input signal is split. The envelope of the input signal is detected with an envelope detector in the top signal path. The envelope is input to a modulator. The output signal of the modulator changes the supply voltage of an amplifier. On the bottom the input signal is fed to a limiter eliminating the envelope information. Essentially, the input signal is disassembled into envelope and phase by envelope detector and limiter, respectively. The modulator is usually a class-S amplifier.

2.5 Linear Amplification with Nonlinear Components (LINC) and Combined Analog Locked Loop Universal Modulator (CALLUM)

Lastly, linear amplification with nonlinear components (LINC) and combined analog locked loop universal modulator (CALLUM) are presented. LINC and CALLUM both linearize a complete transmitter. Fig. 2.9 depicts a basic LINC technique [1], [5]. The input signal is separated into two constant envelope and phase-modulated components by a DSP block, illustrated on the left side in Fig. 2.9. Each signal is up-converted and drives a nonlinear power amplifier. Both PA increase the signals by the same amount. Finally, the outputs of both power amplifiers are summed on the right side in Fig. 2.9. The final output signal is (ideally) an amplified, distortionless version of the input signal [16]. Nonlinear, high efficient power amplifiers are utilized in the LINC

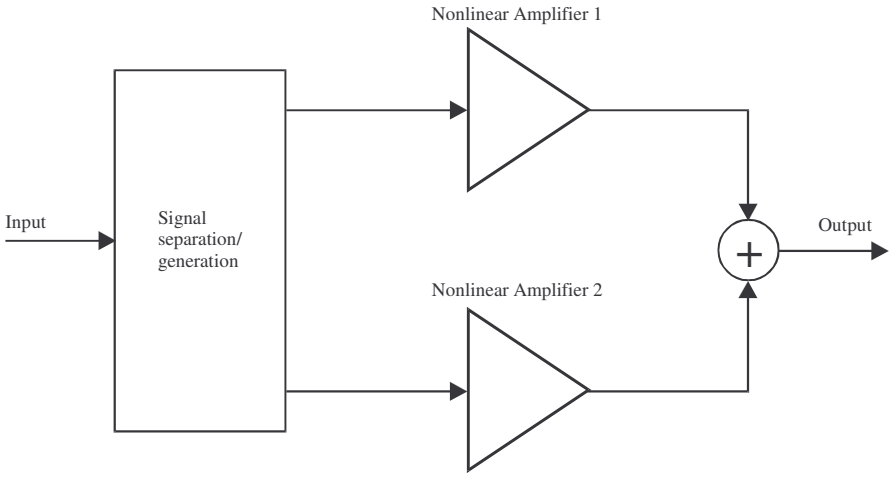

Figure 2.9 Linear amplification using nonlinear components (LINC): the input signal on the left side is input to a signal separation circuit. The output signals are of constant amplitude with 90 degree phase difference and amplified with nonlinear amplifiers. The outputs of the two nonlinear amplifiers are added on the right side, recreating an amplified and linearized input signal. The signal separation network is typically a digital network.

linearization method. The nonlinear power amplifiers create unwanted distortion components with opposite phase at their output. As result, the distortion components cancel each other after summation.

CALLUM [17] is a technique, which is derived from LINC. It uses Cartesian feedback. The output signal is fed back, down-converted to I and Q components by a QAM, and compared with the baseband (I and Q) components.

High efficiency can be achieved with LINC and CALLUM because of the use of efficient power amplifiers. However, LINC has difficulties with gain and phase matching in both paths, thus resulting in imperfect cancellation of distortion. CALLUM suffers from the limitations of a feedback system. It is typically used for narrow-band applications. Finally, power and efficiency is lost because the summation takes place at

the output of the power amplifiers [1].

Other linearization techniques are proposed in [18], [19] and [20]. Yang et al. [18] presents a method where a second order intermodulation term is generated at the input of a main amplifier (gate of a transistor). The second order intermodulation component is then fed forward through an auxiliary feedforward amplifier to the output of the main amplifier (drain of a transistor), resulting in reduction of third order distortion at the final output. However, the proposed method allows only moderate changes in input power.

Ding et al. [19] propose a linearized low noise amplifier (LNA). The circuit is arranged as a differential LNA. Two identical LNA, one LNA is the main amplifier, the second amplifier is considered as the auxiliary amplifier, are connected in parallel. A fraction of the current in the auxiliary amplifier is subtracted from the current in the main amplifier. Intermodulation products are reduced if they are opposite in phase. The input to the auxiliary amplifier is multiplied with a factor β, in order to achieve the desired reduction in third order distortion. However, it is not clearly stated in [19] how β is created. The factor β indicates the presence of an additional ideal amplifier with gain β, at the input of the auxiliary LNA. More disadvantages are, for instance, increased power consumption and gain loss due to subtraction of the linear components.

Webster et al. [20] proposes a method for reduction of third order intermodulation products particularly for use in monolithic microwave integrated circuits. This new method is called derivative superposition technique. In this technique two or more MOSFETs are connected parallel. The nonlinear drain current of a MOS transistor can be expressed as a function of gate-source voltage with a Taylor series expansion. The third coefficient of the series is the third derivative of the transconductance of a MOSFET.

Furthermore, the third derivative of the transconductance is dependent on the gate width and its sign (positive or negative) varies with bias current. If the two transistors are biased, such that the third derivatives of the transconductance have different signs, then a reduction of third order distortion is achieved by subtraction. However, variations in transistors due to manufacturing and temperature limit the linearization performance.

Finally, a method to linearize the DC transfer characteristic of differential amplifiers is presented in [21] This method uses two differential pairs with cross-coupled drains to achieve subtraction at the output of the combined differential pairs, thus the name cross-coupled differential pairs. The approach for reduction of distortion used in [21] employs the fact that the transfer characteristic of differential pairs can be described with a Taylor series expansion. With the proper adjustment of bias currents and aspect ratios of the transistors, cancellation of the third order component of the Taylor series can be reached and third order distortion reduced. However, the approach presented is based on the square-law behavior of MOS transistors. The derived expression for cancellation in [21] is not valid if higher order effects such as velocity saturation in transistors appear. Furthermore, Taylor series expansion is very cumbersome for complex circuit topologies, and does not give informative insight into the design approach of circuits. Conversely, the analytic method presented in the present thesis is general and does not depend on underlying assumptions such as the square-law characteristic of MOS transistors. Finally, the groundwork presented in this thesis is applicable to a wider variety of linearization implementations beyond the simple cross-coupled differential pair. Nevertheless, the cross-coupled differential pair provides simple baseline validation of a specific case of a wider range of approaches discussed in this thesis.

Table 2.1 summarizes the linearization techniques. The linearization techniques proposed in [18], [19] and [20] best fit into the feedforward category. In addition, a comparison of suppression of third order distortion for prior linearization methods is listed in Table 2.2. All data is the result of a two-tone test input signal except where otherwise noted. Adjacent channel power ratio (ACPR) is defined as distortion level in the adjacent channel to the desired channel relative to the power level in the desired channel in dB.

Table 2.1 Prior linearization techniques [12]

Feedback	Feedforward	Predistortion
RF feedback	Basic feedforward	RF/IF predistortion
Cartesian Loop Polar Loop	Envelope Elimination and Restoration	Linear amplification using nonlinear components (LINC)/Combined analog locked loop universal modulator (CALLUM)
Adaptive baseband predistortion	**New linearization method**	Adaptive baseband predistortion

Table 2.2 Summary of third order distortion suppression of prior linearization techniques

Linearization technique	Third order distortion suppression	Reference
RF Feedback	12 dB	[22]
Envelope Feedback	10 dB	[3]
Polar Loop Feedback	30 dB	[3]
Cartesian Loop Feedback	30 dB	[23]
Feedforward	20 – 40 dB 22 dB	[9] [11]
RF/IF Predistortion	10 dB in ACPR 20 dB	[3] [24]
Adaptive Baseband Predistortion	30 dB	[5]
Envelope Elimination & Restoration (EER)	35 dB	[14]
LINC	35 dB	[25]
Second order intermodulation feedforwarding	18 dB	[18]
Beta feedforwarding	15 dB	[19]
Derivative superposition	15 dB @10 dBm output power compared to class AB amplifier	[20]

CHAPTER 3: THEORETICAL BACKGROUND

Before proceeding to describe the new linearization method in Chapter 4, the present chapter describes the concepts of nonlinearity, underlying mathematical models, and the effects of nonlinearity on radio systems. Depending on the operating region (bias and signal level) and the degree of nonlinearity in the input-output transfer characteristic, nonlinear circuits can roughly be characterized as weakly nonlinear or strongly nonlinear [25].

The transfer characteristic of weakly nonlinear and memoryless circuits can be described by a power series expansion such as a Taylor series expansion, where the series is truncated after the first few terms [25]. For strongly nonlinear circuits, the power series becomes unwieldy because the series needs many terms with large coefficients to accurately model the nonlinear behavior. Conversely, weakly nonlinear circuits are more easily modeled with a few terms.

In practice, an amplifier has to operate well below the gain compression (saturation) to be in the weakly nonlinear region. An amplifier is considered strongly nonlinear, if the amplifier operates close to the 1dB compression point.

3.1 Modeling of Nonlinear Behavior

The input-output characteristic of weakly nonlinear circuits is commonly modeled with a power series expansion [2] as shown in (3.1).

$$y(t) = \sum_{n=0}^{\infty} c_n x(t) = c_0 + c_1 x(t) + c_2 x^2(t) + c_3 x^3(t) + \ldots \qquad (3.1)$$

In equation (3.1), x(t) corresponds to a time varying input signal and y(t) to the resulting time varying output signal. Coefficient c_0 corresponds to a DC offset and c_1 to the linear component. Coefficients c_2 and c_3 represent nonlinearities of second and third order, respectively.

For a linear circuit, the series (3.1) would reduce to $y(t)=c_1 x(t)$, where $c_0=c_2=c_3=0$ and $c_1=10$, for example. For a voltage amplifier, where x(t) and y(t) are voltages, coefficient c_1 would be equal to the voltage gain. Fig. 3.1 illustrates the linear behavior of an amplifier in time and frequency domain. In Fig. 3.1(a), the input signal x(t) on the left side is equal to a signal of the form $A\sin(\omega_0 t)$ where A is the maximum amplitude and equal to one and ω_0 is the frequency in radians per second. If $c_1=10$, then the output signal y(t), depicted on the right side in Fig. 3.1(a) equals $10\ A\sin(\omega_0 t)$. Fig. 3.1(b) and (c) depict the frequency spectra of input and output signals, respectively. The input and output spectra only have one frequency, the same frequency at input and output. The amplitude of the output signal is ten times greater than the amplitude of the input signal.

The frequency spectra of the input and output signals can be obtained by applying the Fourier transform. The Fourier transform of a sinusoid, $\cos(\omega_0 t)$ is equals to (3.2) [27].

$$F\{\cos(\omega_0 t)\} = \int_{-\infty}^{\infty} \cos(\omega_0 t) e^{-j\omega t}\, dt = \delta(\omega - \omega_0) + \delta(\omega + \omega_0)\,. \qquad (3.2)$$

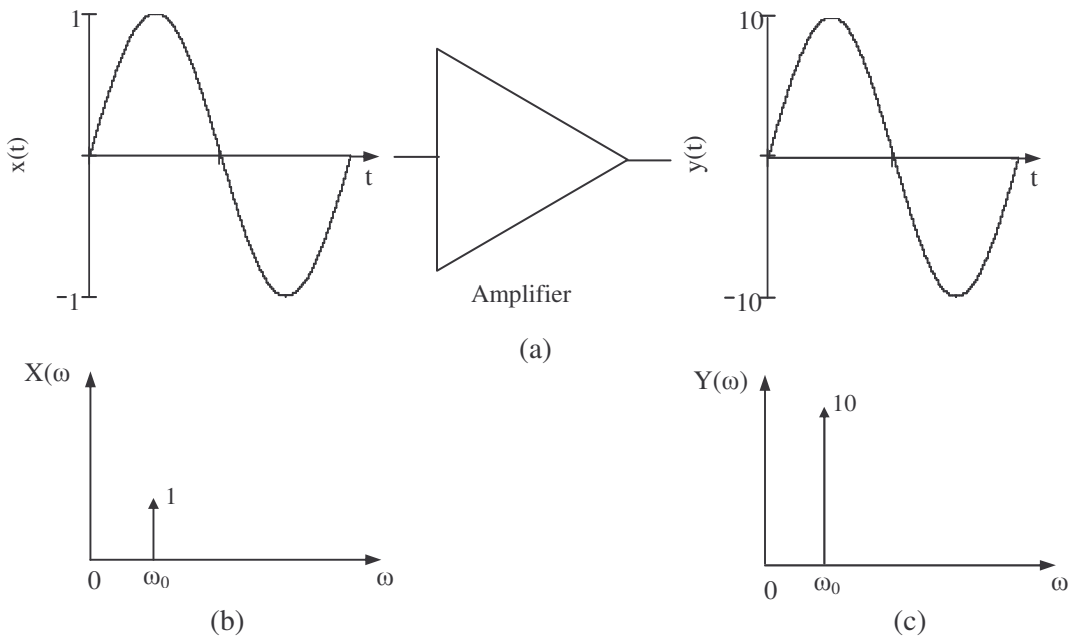

Figure 3.1 Input and output signals of linear amplifier: the left side in (a) shows the input signal x(t) in time domain with amplitude 1. The right side in (a) depicts the output signal y(t) in the time domain. (b) displays the frequency spectrum of the input signal X(ω) with a spectral line at ω_0 and (c) shows the spectrum of the output signal Y(ω) with a spectral line at ω_0. Coefficient c_1=10.

The frequency is denoted ω in (3.2) and δ is the dirac impulse function. Only the positive frequency is shown in Fig. 3.1(b) and (c) for the example of a linear circuit.

For the above case, the output signal y(t) over input x(t) is depicted in Fig. 3.2.

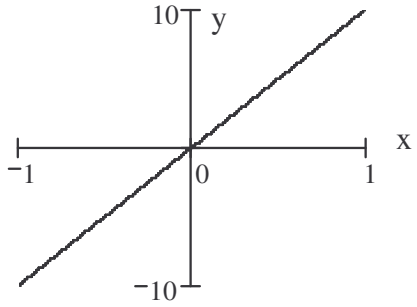

Figure 3.2 Transfer characteristic of linear amplifier; the constant slope equals c_1=10.

The "straight-line" with a slope of c_1=10 displays a linear transfer characteristic.

In cases where the amplifier is not linear and exhibits a second order nonlinearity, the transfer relationship would be equation (3.3) [1]

$$y(t) = c_1 x(t) + c_2 x^2(t), \tag{3.3}$$

where c_0=c_3=0, c_1=10 and c_2=2. Equation (3.3) is presented graphically in Fig. 3.3. The output signal y(t) is shown in Fig. 3.3(a), where the input is x(t)=Asin(ω_0t) and A=1. The output spectrum Y(ω) in Fig. 3.3(b) shows a DC component at ω=0 and a second harmonic at $2\omega_0$. Both components have absolute amplitudes of $|c_2A^2/2|$=1. The

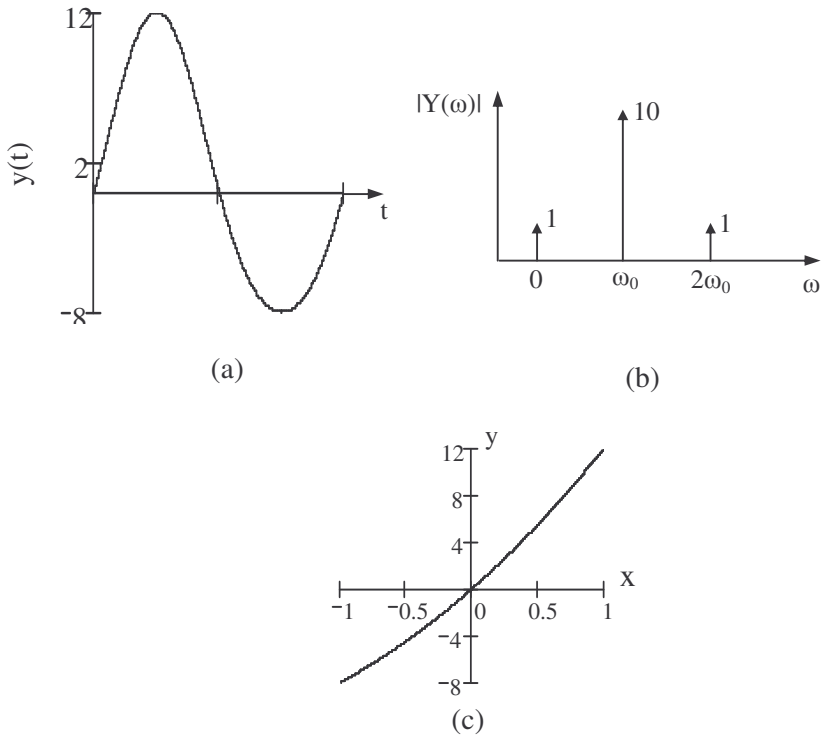

Figure 3.3 Second order nonlinearity in amplifier: (a) Output signal y(t) in time domain, signal exhibits positive DC offset; (b) Output signal |Y(ω)| in frequency domain, DC component at ω=0 and second harmonic at ω=$2\omega_0$ with fundamental at ω_0; (c) input-output transfer characteristic has nonlinear behavior; c_1=10, c_2=2, c_0=c_3=0

fundamental signal appears at ω_0 with an absolute amplitude of $|c_1A|=10$. Fig. 3.3(c) depicts the transfer characteristic. Due to the second order nonlinearity, the transfer relationship deviates from a linear "straight-line" input-output relationship. The full expansion of equation (3.3) is presented in Appendix A.

If the amplifier exhibits a third order nonlinearity, (3.1) can be simplified to (3.4)

$$y(t) = c_1 x(t) + c_3 x^3(t), \tag{3.4}$$

where $c_0=c_2=0$, $c_1=10$ and $c_3=-3$. Fig. 3.4 demonstrates the resultant nonlinear behavior in

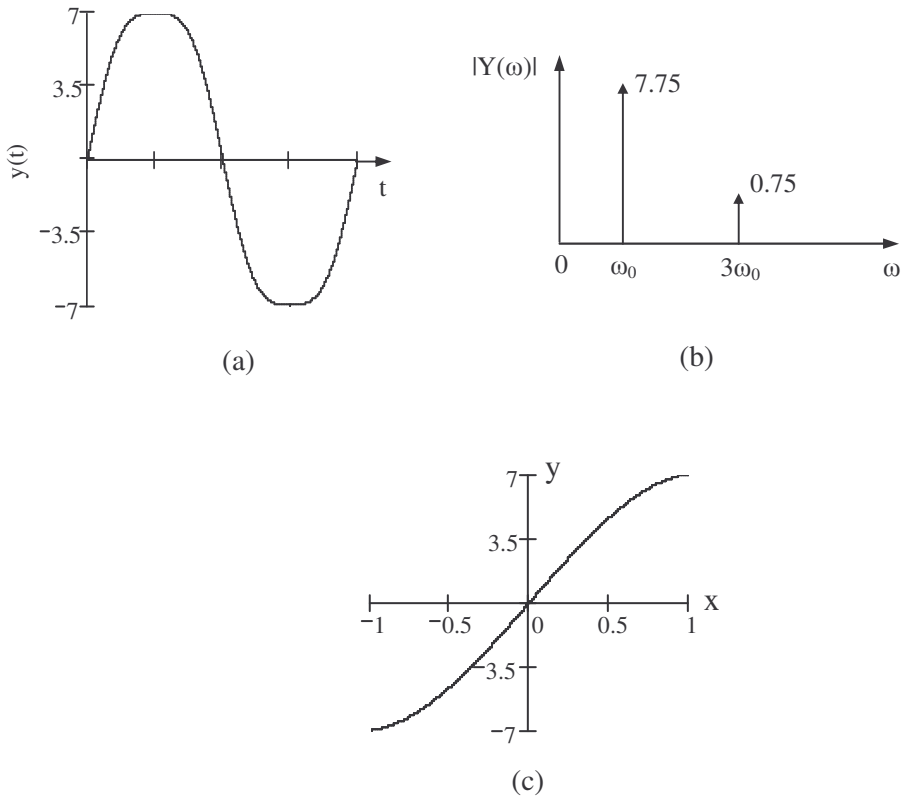

Figure 3.4 Third order nonlinearity in amplifier: (a) Output signal y(t) in time domain; signal exhibits compression; (b) output signal $|Y(\omega)|$ in frequency domain, third order harmonic at $\omega=3\omega_0$ and fundamental at ω_0; (c) input-output transfer characteristic of amplifier with third order nonlinearity deviates from linear characteristic; $c_1=10$, $c_0=c_2=0$, $c_3=-3$

time and frequency domain, where $x(t)=A\sin(\omega_0 t)$ and A=1. The output signal y(t) is displayed in the time domain in Fig. 3.4(a). The clipping of output waveform y(t) in the time domain is an effect of third order nonlinearities. The output spectrum $Y(\omega)$ in Fig. 3.4(b) exhibits a third harmonic at $3\omega_0$, with an absolute amplitude of $|c_3A^3/4|=0.75$. The spectrum $Y(\omega)$ does not show a DC component. Furthermore, the fundamental signal at ω_0 has an absolute amplitude of $|(c_1A+3c_3A^3/4)|=7.75$. Finally, the transfer characteristic of equation (3.4) is depicted in Fig. 3.4(c) where the clipping effect is also visible for higher input signal levels. Again, a full expansion of equation (3.4) is presented in Appendix A.

The fourth case combines second and third order nonlinearities, thus creating second and third harmonics as shown in Fig. 3.5. The describing equation (3.1) is altered to equation (3.5).

$$y(t) = c_1 x(t) + c_2 x^2(t) + c_3 x^3(t) \tag{3.5}$$

In equation (3.5), $c_0=0$, $c_1=10$, $c_2=2$ and $c_3=-3$. The output signal y(t) in the time domain is illustrated in Fig. 3.5(a), with an input waveform $x(t)=A\sin(\omega_0 t)$, where A=1. Output y(t) exhibits a positive DC offset. Output spectrum $Y(\omega)$, depicted in Fig.3.5(b) contains a DC component at $\omega=0$ as well. Furthermore, second and third order components in equation (3.5) create second harmonic at $2\omega_0$, and third harmonic at $3\omega_0$ in the output spectrum $Y(\omega)$ in Fig 3.5(b). The input-output transfer characteristic, which deviates from a linear transfer characteristic, is displayed in Fig. 3.5(c). Appendix A presents a full expansion of equation (3.5) with $x(t)=A\sin(\omega_0 t)$. The linear, second and third order terms, resulting from the full expansion of equation (3.5), are summarized in Table 3.1.

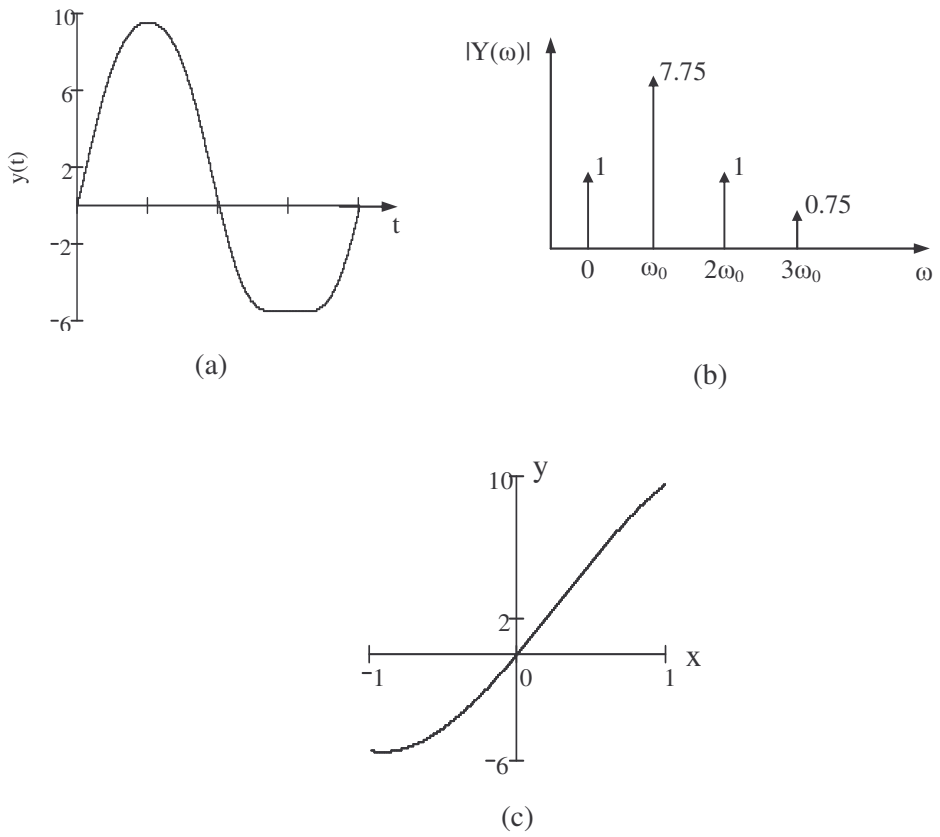

Figure 3.5 Second and third order nonlinearities in amplifier: (a) output signal y(t) in time domain showing DC offset and compression; (b) frequency spectrum of output signal $|Y(\omega)|$; fundamental ω_0, second harmonic $2\omega_0$ and third harmonic $3\omega_0$ as well as a DC component at $\omega=0$ appear in the output spectrum; (c) depicts the nonlinear input-output transfer characteristic. $c_0=0$, $c_1=10$, $c_2=2$, $c_3=-3$

In Table 3.1, A is the input signal amplitude and Coefficients c_1, c_2 and c_3 are linear, second and third order coefficients of equation (3.5), respectively.

The previous discussion showed various terms of the power series expansion. All of these terms are strongly dependent on the coefficients c_1, c_2 and c_3. The coefficients are usually not easily available. Furthermore, the accuracy of the coefficients is dependent on biasing and technology parameters, and determined by device models such as transistor models. Particularly for transistor models, the accuracy of the coefficients

Table 3.1 Summary of linear and higher order terms of power series expansion

DC term	$\frac{1}{2}c_2A^2$
First order term at fundamental frequency	$c_1A+\frac{3}{4}c_3A^3$
Second order term at second harmonic	$\frac{1}{2}c_2A^2$
Third order term at third harmonic	$\frac{1}{4}c_3A^3$

depends on the inclusion of second order effects in transistors, for instance velocity saturation and mobility reduction.

3.2 Effects of Nonlinearity in Radio Systems

In the following sections, a variety of nonlinear effects in radio systems are discussed. First, two-tone test as a means of characterizing nonlinearity is presented. Secondly, the nonlinear effects desensitization, cross modulation and intermodulation are discussed.

3.2.1 Two-Tone Test

In many radio systems, an amplitude modulation scheme is used to transmit information. To exercise signal amplitude, a two-tone test provides a simple time varying signal envelope in order to test a circuit such as an amplifier.

In the two-tone test, the input test signal consists of two sinusoids with frequencies closely spaced in the band of interest as shown in (3.6).

$$x(t) = A_1\cos(\omega_1 t) + A_2\cos(\omega_2 t) \tag{3.6}$$

In equation (3.6), A_1 and A_2 are the amplitudes of the sinusoids and ω_1 and ω_2 are the frequencies in radians per second. Usually, both sinusoids have equal amplitude $A_1=A_2$.

Fig. 3.6(a) illustrates a two-tone signal in the time domain with $A_1=A_2=1$. The peak amplitude is $2A_1$ and equals twice the amplitude of each sinusoid. The envelope changes with the difference frequency ω_1-ω_2. The corresponding frequency spectrum in Fig. 3.6(b) consists of the two frequencies ω_1 and ω_2 with equal amplitudes.

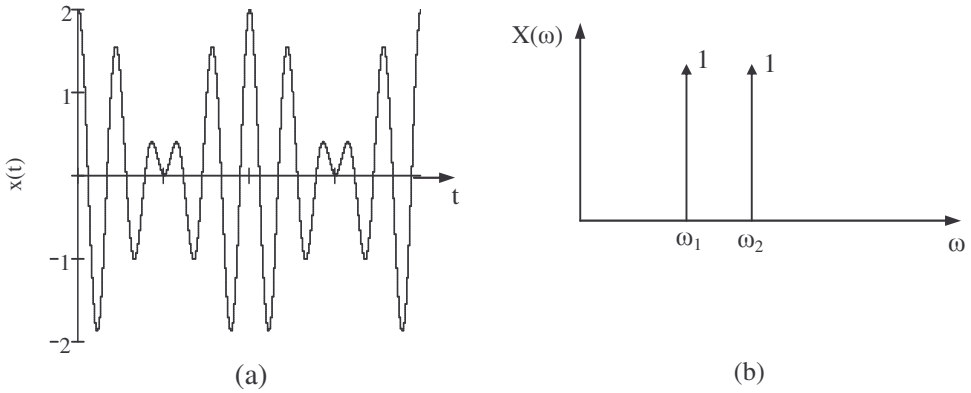

Figure 3.6 Two-Tone signal: (a) time domain with varying envelope from 0 to $2A_1$; (b) frequency domain showing two sinusoids at frequencies ω_1 and ω_2.

If an amplifier exhibits a nonlinear transfer characteristic such as (3.5) and is excited with a two-tone signal (3.6), the output signal will contain second and third order harmonics and intermodulation products. The output signal y(t) then is

$$\begin{aligned} y(t) &= c_1 x(t) + c_2 x^2(t) + c_3 x^3(t) \\ &= c_1(A_1\cos(\omega_1 t) + A_2\cos(\omega_2 t)) + c_2(A_1\cos(\omega_1 t) + A_2\cos(\omega_2 t))^2 + c_3(A_1\cos(\omega_1 t) + A_2\cos(\omega_2 t))^3 \end{aligned} \tag{3.7}$$

With the use of common trigonometric identities, (3.7) can be expanded as shown in [27] and Appendix A. A few important terms are used to obtain key results in the following

discussions.

3.2.2 Desensitization

In radio receivers, one effect of nonlinearity is when weak signals are being blocked by strong signals at different frequencies. This situation can occur when a radio receiver has to process a weak signal radiated from a remote radio sender in the presence of a strong signal emitted from a close radio transmitter. When this strong signal is processed along with the weak desired signal, a radio system can exhibit an effect called desensitization [2], [28]. This effect can be explained with the fundamental term of the full expansion of equation (3.7), which is described with (3.8).

$$(c_1 A_1 + \frac{3}{4} c_3 A_1^3 + \frac{3}{2} c_3 A_1 A_2^2) \cos(\omega_1 t) \tag{3.8}$$

where (3.8) represents the output sinusoid at the same frequency as the input, i.e. (3.8) represents the linear component of the output signal. Coefficients c_1 and c_3 represent the linear component and third order nonlinearity of the nonlinear amplifier, respectively. If A_1 is the amplitude of the weak desired signal in (3.8) and A_2 the amplitude of a strong interferer and $A_1 << A_2$, (3.8) can be approximated as shown in (3.9), where the term $c_1+3c_3A_2^2/3$ represents the linear gain of the nonlinear amplifier.

$$(c_1 + \frac{3}{2} c_3 A_2^2) A_1 \cos(\omega_1 t) . \tag{3.9}$$

For $c_3<0$, the gain as given by $c_1+3c_3A_2^2/3$, is a decreasing function of A_2 [2]. This decrease in gain decreases the desired signal strength at the output of the amplifier.

3.2.3 Cross Modulation

A second effect of nonlinearity in radio systems is cross modulation. Cross modulation is the nonlinear effect where modulation from one carrier is transferred to another [2], [28]. To illustrate cross modulation, consider an input test signal where the test signal is an amplitude-modulated carrier at frequency ω_2 added to a sinusoidal waveform at frequency ω_1. The resulting signal can be described with equation (3.10):

$$x(t) = A_1 \cos(\omega_1 t) + A_2 (1 + m\cos(\omega_m t))\cos(\omega_2 t)\,. \tag{3.10}$$

In equation (3.10), A_1 and A_2 are the peak amplitudes of the carriers with frequencies ω_1 and ω_2 in radians per second, m is the modulation index ($0 \le m \le 1$), and ω_m is the modulating frequency in radians per second. If amplitude A_2 in the fundamental component of the full expansion of equation (3.7) is replaced by $A_2(1+m\cos(\omega_m t))$, then the fundamental component of the full expansion of (3.7) becomes (3.11)

$$(c_1 A_1 + \frac{3}{4} c_3 A_1^3 + \frac{3}{2} c_3 A_1 A_2^2 (1 + m\cos(\omega_m t))^2)\cos(\omega_1 t)\,. \tag{3.11}$$

As can be seen from (3.11), the amplitude of carrier ω_1 now includes a time-varying amplitude modulation term $(1+m\cos(\omega_m t))^2$ that induces modulation on the unmodulated carrier at frequency ω_1. Cross modulation is particularly critical in multi carrier radio systems.

3.2.4 Intermodulation

A third effect of nonlinearity on radio circuits is intermodulation. In intermodulation, a nonlinear circuit excited with two input sinusoids results in an output,

containing components at new frequencies. In the general case, these new frequencies will be of the form as described with equation (3.12) [29].

$$\omega = m\omega_1 \pm n\omega_2 . \tag{3.12}$$

In equation (3.12), ω is the new frequency; ω_1 and ω_2 are the original frequencies in radians per second. m and n are positive integers including zero. m+n is equal to the order of distortion.

Not all new created frequencies belong to the group of intermodulation products. Intermodulation products are the result of the interaction of fundamentals and harmonics of different order. Harmonics are typically not referred to as intermodulation products.

Intermodulation products arise from the $x^2(t)$, $x^3(t)$ and higher order terms of the power series in (3.1). Products created by the $x^2(t)$ term are second order products and products created by the term $x^3(t)$ are third order products.

Returning to the two-tone input signal of equation (3.6), it can be seen from the full expansion of (3.7) that the following intermodulation products appear:

Second order intermodulation products:

$$c_2 A_1 A_2 \cos((\omega_1 + \omega_2)t) + c_2 A_1 A_2 \cos((\omega_1 - \omega_2)t) \tag{3.13}$$

For $A_1=A_2$, term (3.13) changes to term (3.14):

$$c_2 A_1^{\,2} \cos((\omega_1 + \omega_2)t) + c_2 A_1^{\,2} \cos((\omega_1 - \omega_2)t) \tag{3.14}$$

In terms (3.13) and (3.14), A_1 and A_2 are the amplitudes of the two-tone input signal and c_2 is the second order coefficient. ω_1 and ω_2 are the fundamental frequencies in radians per second.

Third order intermodulation products:

$$\frac{3}{4}c_3A_1^2A_2\cos((2\omega_1+\omega_2)t)+\frac{3}{4}c_3A_1^2A_2\cos((2\omega_1-\omega_2)t)$$
$$+\frac{3}{4}c_3A_1A_2^2\cos((2\omega_2+\omega_1)t)+\frac{3}{4}c_3A_1A_2^2\cos((2\omega_2-\omega_1)t) \quad (3.15)$$

Again, for $A_1=A_2$ terms(3.15) change to terms (3.26):

$$\frac{3}{4}c_3A_1^3\cos((2\omega_1+\omega_2)t)+\frac{3}{4}c_3A_1^3\cos((2\omega_1-\omega_2)t)$$
$$+\frac{3}{4}c_3A_1^3\cos((2\omega_2+\omega_1)t)+\frac{3}{4}c_3A_1^3\cos((2\omega_2-\omega_1)t) \quad (3.16)$$

In terms (3.16), A_1 and A_2 are the amplitudes, c_3 is the third order coefficient of the power series, and ω_1 and ω_2 are the two input frequencies in radians per second. In terms (3.13) to (3.16) harmonics are not considered.

Term (3.16) indicates that as input amplitude A1 increases by a factor of two, the third order intermodulation product goes up by a factor of eight due to the exponent of value three in the terms $3c_3A_1^3/4$. A factor of two in voltage equals 6 dB and a factor of eight equals 18 dB.

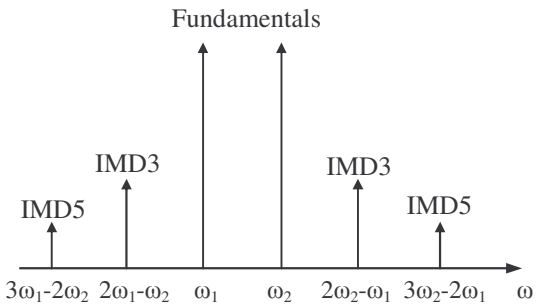

Figure 3.7 Frequency spectrum of intermodulation distortion: fundamental tones at frequencies ω_1 and ω_2, third order intermodulation distortion (IMD3) products at frequencies $2\omega_1-\omega_2$ and $2\omega_2-\omega_1$, and fifth order intermodulation distortion (IMD5) products at frequencies $3\omega_1-2\omega_2$ and $3\omega_2-2\omega_1$

Of great interest are intermodulation products appearing in the vicinity of the carrier frequency. These are third order intermodulation products of the form $2\omega_1-\omega_2$ and $2\omega_2-\omega_1$. Fig. 3.7 depicts the two tones with the intermodulation distortion (IMD) of interest. The two fundamental tones of a two-tone signal are the two innermost lines in Fig. 3.7. Third order intermodulation products appear at frequencies $2\omega_1-\omega_2$ and $2\omega_2-\omega_1$. The outermost tones at frequencies $3\omega_1-2\omega_2$ and $3\omega_2-2\omega_1$ are fifth order intermodulation products.

3.3 Measures of Nonlinearity

In order to get a better understanding of the level of nonlinear distortion of a circuit, measures are defined which characterize the circuits' nonlinear behavior. Two commonly used measures to describe nonlinearities of a circuit are 1dB compression point (P1dB) and third order intercept point (IP3). Simple test signals such as those described in the previous section are commonly used to test radio circuits for linearity and find these measures. First, the 1dB compression point as a measure for gain

compression will be discussed. The derivation of the intercept point will be shown next.

3.3.1 1dB Compression Point

The 1dB compression point is defined as the power level where the small signal gain is degraded by 1 dB. P1dB can be input (IP1dB) or output referred (OP1dB) and is usually given in units of dBm. Typically, the output compression point is used, and will be described in the following section. To measure the 1dB compression point, a single sinusoidal test signal y(t)=Acos(ωt) is usually supplied to a nonlinear circuit with A as the peak amplitude and ω as the frequency in radians per second. If the circuit under test has the nonlinear transfer characteristic of (3.1) up to third order, the resulting output signal is of the form presented with equation (3.17).

$$y(t) = \frac{1}{2}c_2A^2 + (c_1A + \frac{3}{4}c_3A^3)\cos(\omega t) + \frac{1}{2}c_2A^2\cos(2\omega t) + \frac{1}{4}c_3A^3\cos(3\omega t) \qquad (3.17)$$

Coefficients c_1, c_2 and c_3 represent the first, second and third order nonlinearities of the circuit under test, respectively. The term $c_1A+3c_3A^3/4$ represents the small signal gain at the fundamental frequency. Assuming that c_1A is greater than the other terms containing A, the small signal gain equals c_1 for small input signals. If $c_3<0$, the gain approaches zero for sufficiently high input amplitude levels [2].

The 1dB compression point is measured by increasing the input amplitude A from zero until the measured gain decreases by 1 dB. A compression of 1 dB (decrease in gain by 1 dB) corresponds to an ideal voltage gain degradation by a factor of $10^{-0.05} = 0.89125$ [2]. Since the linear uncompressed gain is c_1, 1dB compression then occurs as described with equation (3.18).

$$(0.89125)c_1 = c_1 + \frac{3}{4}c_3 A^2 . \tag{3.18}$$

In equation (3.18), A is the input signal amplitude, c_1 is the linear coefficient and c_3 is the third order coefficient. Solving for the input signal amplitude A where 1dB compression occurs for $c_3<0$ results in equation (3.19), where A_{1dBin} is the input signal level when 1 dB compressed.

$$A_{1dBin} = 0.3808\sqrt{\left|\frac{c_1}{c_3}\right|} \tag{3.19}$$

The output signal amplitude A_{1dBout}, when 1 dB compressed, can be found by setting $0.89125c_1A_{1dBin}=A_{1dBout}$. The gain equals $c_1 0.89125$ when compressed by 1dB. Substituting A_{1dBin} with (3.19) results in $A_{1dBout}=0.3427 \bullet c_1(|c_1/c_3|)^{0.5}$.

Conversely, one can solve for c_3, given A_{1dBin}. Equation (3.20) is describing coefficient c_3.

$$|c_3| = 0.145\frac{|c_1|}{A_{1dBin}{}^2} \tag{3.20}$$

Again, c_1 and c_3 are the first and third order coefficients of the power series describing the transfer characteristic of a nonlinear device. Fig. 3.8 illustrates the 1dB compression point by displaying output power over input power in logarithmic scale. The dashed line in Fig. 3.8 depicts the linear gain without compression and the solid curve is the compressed gain. The input power level where the two gain curves differ by 1 dB is

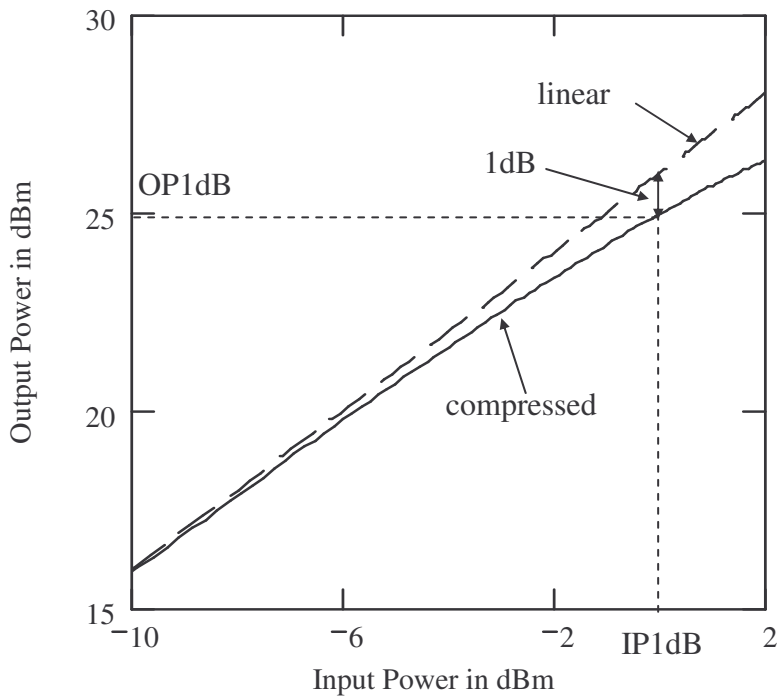

Figure 3.8 1dB-compression point: linear transfer (dashed line) and compressed transfer characteristic (solid line); input referred 1dB compression point, IP1dB is the input power level where the gain is compressed by 1 dB, the compressed output power level at this input power level is called output referred 1dB compression point OP1dB

called 1dB input compression point. Conversely, the corresponding output power level of the compressed gain curve is denoted 1dB output compression point.

3.3.2 Intercept Point

Intermodulation products, particularly third order distortion products, are of major concern in communication systems. As discussed above, the frequencies of these products appear close to the desired band and are difficult to filter. The level of intermodulation products is dependent on input power level. The intercept point (IP) is a single parameter that characterizes the behavior of intermodulation products independent

of input power level [1]. The intercept point can be input or output referred. Of most interest are second order intercept point (IP2) and third order intercept point (IP3).

The intercept point is defined as the point where the fundamental linear component and intermodulation products have equal amplitude at the output of a nonlinear circuit. In most practical circuits, intermodulation products will never be equal to the fundamental linear term because both amplitudes will compress before reaching this point. Nevertheless, the intercept point is useful to characterize circuits.

In the following section the third order intercept point will be derived, using a two-tone test signal with equal amplitudes $A=A_1=A_2$ as input to a nonlinear circuit. At the theoretical intercept point, the uncompressed linear component c_1A would equal the third order intermodulation product from equation (3.16), as follows from equation (3.7) and results in equation (3.21).

$$c_1 A = \frac{3}{4} c_3 A^3 \tag{3.21}$$

Amplitude A in equation (3.21) is the input signal amplitude of the two-tone test, c_1 is the small signal gain and c_3 represents the third order coefficient. Solving for $A=A_{IIP3}$ gives

$$A_{IIP3} = \sqrt{\frac{4}{3}\left|\frac{c_1}{c_3}\right|} \tag{3.22}$$

Amplitude A_{IIP3} in equation (3.22) is the input signal level where the linear component is equal to the third order intermodulation product. Coefficients c_1 and c_3 describe linear and third order nonlinearities.

Typically, power levels are of interest in radio systems. The input power level

with unit Watts is found by dividing $(A_{IIP3})^2$ by two times the input resistance R_{in}, if A_{IIP3} is a voltage level [29]. The resulting input power is described with expression (3.23), where P_{IIP3} is the input power level in Watts, c_1 is the small signal gain and c_3 is the third order coefficient.

$$P_{IIP3} = \frac{2}{3}\left|\frac{c_1}{c_3}\right|\frac{1}{R_{in}} \tag{3.23}$$

The third order input intercept point in unit dBm is found using the following equation.

$$IIP3 = 10\log_{10}\left(\frac{P_{IIP3}}{0.001W}\right)dBm \tag{3.24}$$

In (3.24), IIP3 is the input referred IP3 (IIP3) [29]. In equation (3.24), IIP3 is calculated in respect to 1 mW.

The output amplitude A_{OIP3} is equal to c_1 A_{IIP3} as shown in equation (3.25), where c_1 is the linear and c_3 is the third order coefficient.

$$A_{OIP3} = c_1 A_{IIP3} = \sqrt{\frac{4}{3}\left|\frac{c_1^3}{c_3}\right|}, \tag{3.25}$$

Similar to the input power level, the output power level is calculated by dividing $(A_{OIP3})^2$ by $2R_L$ where R_L is the load resistance and A_{OIP3} is a voltage amplitude. This is shown with expression (3.26).

$$P_{OIP3} = \frac{2}{3}\frac{{c_1}^3}{c_3}\frac{1}{R_L}. \tag{3.26}$$

As stated above, c_1 and c_3 are linear and third order coefficients, respectively. The output referred IP3 (OIP3) in dBm is given by the following equation (3.27), where P_{OIP3} is the output power level in Watts.

$$OIP3 = 10\log_{10}\left(\frac{P_{OIP3}}{0.001W}\right)dBm, \tag{3.27}$$

Equation (3.26) indicates that OIP3 depends on the linear coefficient c_1, the coefficient of third order nonlinearity c_3, and R_L. If c_3 approaches zero then P_{OIP3} approaches infinity as expected since the cubic term in (3.1) is multiplied by c_3. Input and output intercept points are related by the gain G in dB: OIP3=IIP3+G.

Fig. 3.9 illustrates the circumstances of the third order intercept point graphically. The plot in Fig. 3.9 displays the output power versus the input power in dBm. The upper solid line describes the linear component and the lower solid line describes the third order distortion product dependent on input power levels. Both solid lines compress and do not increase further due to compression and the limit in supply voltage. The dashed lines in Fig. 3.9, indicate extrapolation of the solid lines. The point where the dashed lines intersect is called third order intercept point. The horizontal coordinate of the intercept point is usually referred to as third order input intercept point IIP3 and the vertical coordinate is the third order output intercept point OIP3. As stated before, as the input power increases by 1 dB, the linear output power also increases by 1 dB, but the third order distortion increases by 3 dB due to a slope of three of the third order distortion component on a logarithmic scale. In Fig. 3.9, the input IP3 equals 15 dBm and the output IP3 is 26 dBm.

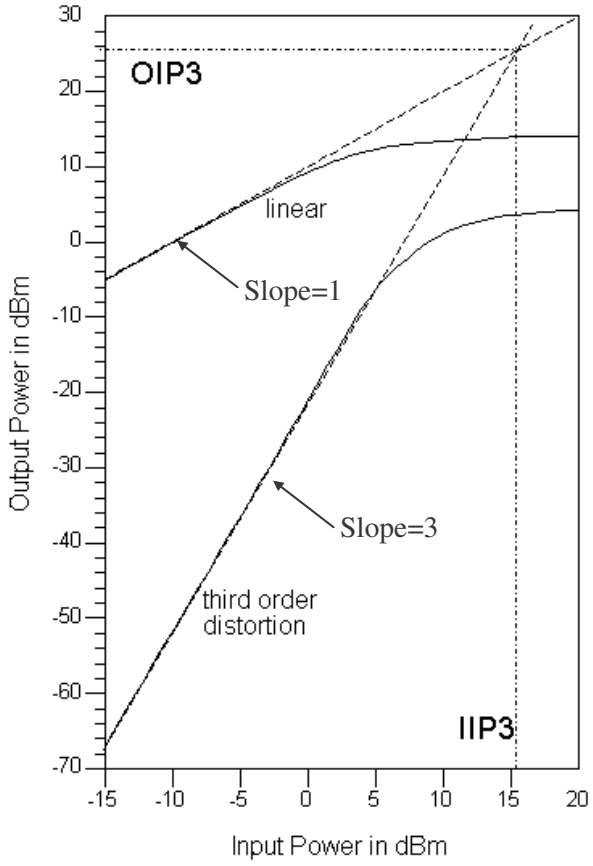

Figure 3.9 Third order intercept point IP3: the linear component has a slope of 1 and the third order distortion has a slope of 3. Solid lines indicate the behavior of a practical circuit, where the output power levels clip. The dashed lines indicate extrapolation of the compressed curves and intercept at the intercept point. IP3 can be input (IIP3=15 dBm) or output (OIP3=26 dBm) referred.

From measurement of the output spectrum using power levels, a more practical equation for calculation of OIP3 can be found with expression (3.28) [29].

$$OIP3 = \frac{P_{out} - P_{3rd}}{2} + P_{out}, \tag{3.28}$$

In the expression (3.28) for the OIP3, P_{out} is the fundamental (linear) power level in dBm, P_{3rd} is the third order distortion product power level in dBm and OIP3 is the third order output intercept point in dBm. Fig. 3.10 illustrates the derivation of equation (3.28). Fig. 3.10 shows the output frequency spectrum of a nonlinear circuit with a two-tone test signal as input. The innermost spectral lines are the fundamental tones of the two-tone test at frequencies ω_1 and ω_2 and the outermost spectral lines are third order distortion products. The linear power levels at the fundamental frequencies ω_1 and ω_2 have equal

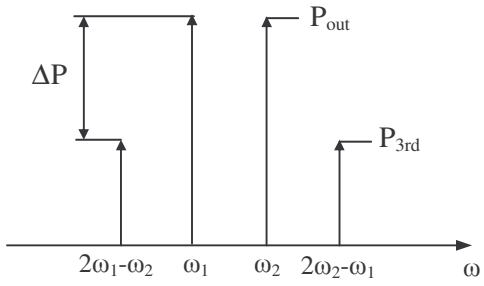

Figure 3.10 Derivation of third order intercept point: fundamental tones at ω_1 and ω_2 have a linear output power level of P_{out} and third order distortion have a power level of P_{3rd}; third order distortion P_{3rd} rises three times faster with increase in input power level than fundamentals, ΔP equals $P_{out}–P_{3rd}=2(OIP3–P_{out})$

output power levels of P_{out} and the third order intermodulation products at frequencies $2\omega_1- \omega_2$ and $2\omega_2-\omega_1$ have an output power level of P_{3rd} in dBm. The difference ΔP equals $P_{out}–P_{3rd}$. P_{out} is the output power level of the fundamental frequency, and P_{3rd} is equal to the power level of the third order intermodulation product. OIP3 is the power level of the third order output intercept point.

A similar intercept point can be defined for the second order intermodulation

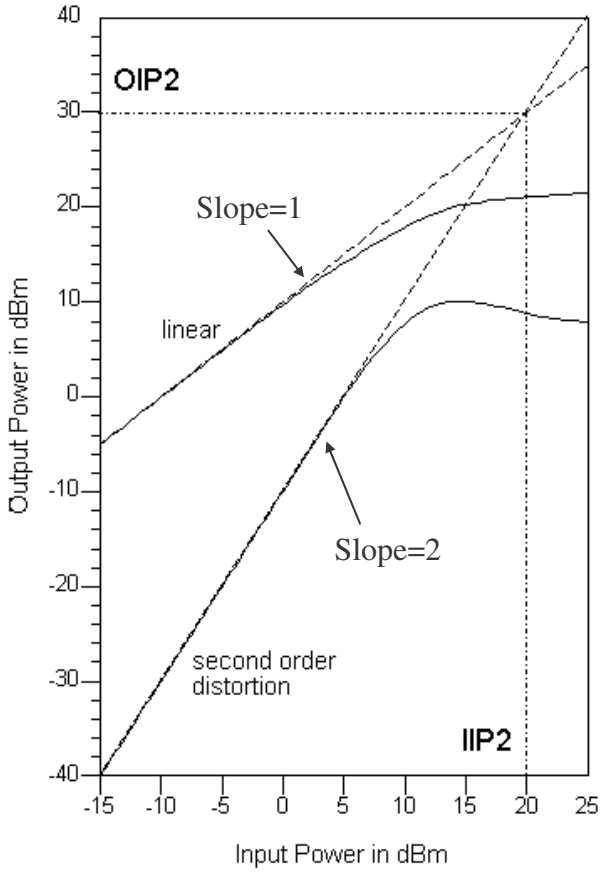

Figure 3.11 Second order intercept point IP2: linear component has slope of one, second order distortion product has a slope of 2; solid lines indicate the behavior of a practical circuit where the output power levels clip. The dashed lines indicate extrapolation of the compressed curves and intercept at the intercept point. IP2 can be input (IIP2=20 dBm) and output (OIP2=30 dBm) referred.

product [28], described with expression (3.29).

$$OIP2 = P_{out} - P_{2nd} + P_{out} \tag{3.29}$$

OIP2 is the second order intercept point in dBm, P_{out} is the linear output power in dBm and P_{2nd} is the second order intermodulation product power level in dBm. Fig. 3.11

illustrates second order intercept point IP2. Similar to Fig. 3.9, the solid lines in Fig. 3.11 are linear and second order distortion components. Both solid curves are extrapolated as shown with the dashed lines. The point where the dashed lines intersect is called second order intercept point. As demonstrated before, the fundamental component increases linearly until clipping occurs. The second order intermodulation product level rises with a slope of two due to the second order term in equation (3.1). IP2 can be referred to input (IIP2) or output (OIP2).

3.4 Phase Distortion

Section 3.1 considers only distortion resulting from changes in input amplitude, commonly called amplitude-modulation-to-amplitude-modulation (AM/AM) conversion. Most practical circuits exhibit memory effects because of storage elements such as capacitors and inductors. In this case, phase distortion can also arise from capacitive or inductive nonlinearities. This is commonly referred to as amplitude-modulation-to-phase-modulation (AM/PM) conversion. It is especially critical in systems where signals with phase modulation, such as quadrature phase shift keying (QPSK), are processed.

The mechanism leading to AM/PM conversion can be explained using the fundamental linear component from (3.17) as shown in (3.30) [26].

$$(c_1 A + \frac{3}{4} c_3 A^3) \cos(\omega t) \tag{3.30}$$

The input signal amplitude is denoted A. c_1 and c_3 are the first and third order coefficients, respectively. The amplitude of equation (3.30) consists of a first order, c_1A, and third order, $3c_3A^3/4$, component. Because of a nonlinear capacitance, a phase

difference can exist between these two components. The response is then the vector sum of two phasors depicted as expression (3.31) [26].

$$V = c_1 A + \frac{3}{4} c_3 A^3 e^{j\theta} \tag{3.31}$$

The sum of the phasors is denoted V and the difference in angle of linear and third order term in (3.31) is denoted θ. c_1 and c_3, again, are first and third order coefficients, respectively. In (3.31), the third order component $3c_3A^3/4$ exhibits a phase angle different from the linear component. Addition of the two phasors, linear and third order, leads to a phasor V with a new angle. This new angle is dependent on variations of the input signal amplitude A, even when θ remains constant [26]. As a result, changes in input amplitude give rise to changes in angle of the linear component of the output signal. This mechanism is called AM/PM conversion. Since the third order component rises three times faster than the first order output component on a logarithmic scale, AM/PM conversion will be most serious as the circuit is driven into saturation.

CHAPTER 4: DESCRIPTION OF NEW LINEARIZATION METHODS

Given the theoretical background of the previous chapter, the new linearization methods are described in the present chapter. The present chapter also presents a simple algebraic expression, useful in practice, which describes the condition for cancellation in terms of output intercept point and gain. The simplicity of the underlying theoretical approach is particularly useful for the design of a wide variety of new linearization topologies, beyond the simple topology built in this thesis. In the following, basic principles of the new linearization approach are presented. Then, conditions for canceling third order distortion are given for a simple embodiment. Finally, an evaluation of the condition of cancellation is performed with Mathcad. Test data for a simple integrated circuit implementation is given later in Chapter 5. More sophisticated linearization embodiments are presented in Chapter 7.

4.1 Basic Principle of New Methods

The underlying principle of the new linearization methods is that a highly distorted signal can be subtracted from a less distorted signal to eliminate the distortion in the final output signal [30]. This overall distortion improvement is achieved by combining a more linear amplifier with a less linear amplifier in such a way that the combined performance is significantly improved and exceeds the performance of the two original amplifiers. This unexpected result can be illustrated with a simple architecture, as shown in Fig. 4.1. In Fig. 4.1, the input signal x is applied to an input coupler which splits

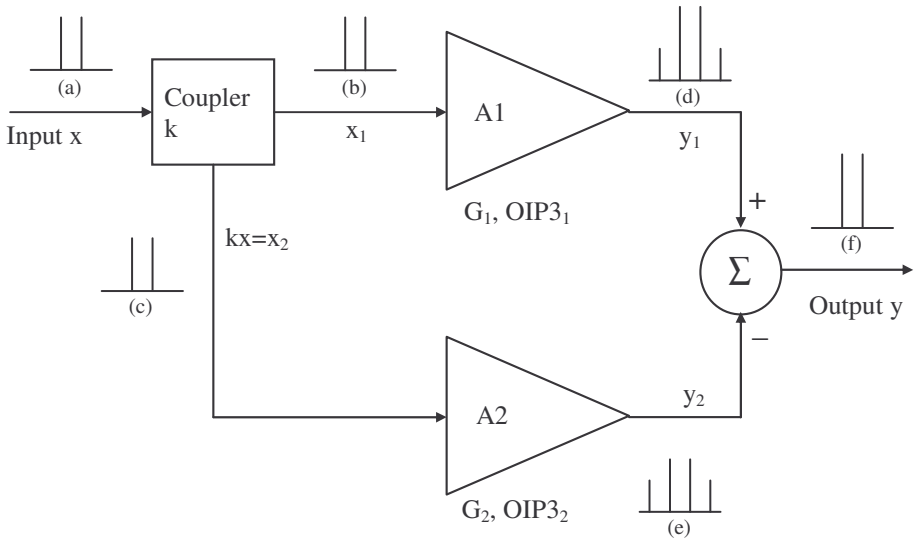

Figure 4.1 Block diagram of new linearization method: coupler splits and scales incoming signal x_1 and x_2, x_1 is input to more linear amplifier A1, x_2 is input to less linear amplifier A2; outputs are subtracted y_1-y_2; frequency spectrum in every path is shown in (a) to (f); gain G and third order output intercept point determine characteristic of amplifier [30].

the signal into x_1 and x_2 shown on the left side in Fig. 4.1. The coupler has a coupling factor k and allows signal x_2, to be scaled relative to x_1. The definition of coupling coefficient k is: $k = x_2/x$. Signal x_1 is then input to more linear amplifier A1 depicted in the upper half of Fig. 4.1. Signal x_2 is input to less linear amplifier A2 illustrated in the lower half of Fig. 4.1. Output y_2 of amplifier A2 is subtracted from the output signal y_1 of amplifier A1 by means of a subtractor as shown on the right side in Fig. 4.1, resulting in final output y. The linearity of the two amplifiers is adjusted by parameters gain G_1 (dB) and output intercept point $OIP3_1$ (dBm) for amplifier A1 and by parameters gain G_2 (dB) and output intercept point $OIP3_2$ (dBm) for amplifier A2. The proper relation of the parameters for cancellation of third order distortion will be derived in Section 4.3.

In Fig. 4.1, the input test signal is a two-tone test signal of the form of equation (3.6). The frequency spectra of all signal paths are shown in Fig. 4.1 labeled from (a) to

(f). First, frequency spectrum (a) is that of a two-tone test signal at the input of the coupler. Frequency spectra (b) and (c) at the inputs of amplifiers A1 and A2, respectively, show signals at the same frequency tones as spectrum (a), when the coupler does not add any distortion.

In Fig. 4.1, the spectrum at the output y_1 of the more linear amplifier A1 is shown in (d). The two outermost lines of spectrum (d) are third order distortion products, whereas the two innermost spectral lines correspond to the linear component. Third order intermodulation products (outermost lines) appear due to nonlinearities in amplifier A1.

Similarly, third order distortion is created at the output of amplifier A2, shown in spectrum (e). In spectrum (e), the fundamental tones have lower amplitude relative to the third order distortion level than the spectrum (d) shown at the output of amplifier A1.

If the parameters of the amplifiers in Fig. 4.1 are adjusted properly, the third order distortion products of (d) and (e) (outermost spectral lines) have the same amplitude. Under such conditions, the final output spectrum shown in (f) only contains the fundamental tones, after subtraction of signal y_2 from y_1. The third order distortion is ideally eliminated. The fundamental tones in (f) are not canceled because fundamental tones in (d) and (e) are different in amplitude. Furthermore, the gains of the two amplifiers are not equal. Finally, the output signal y only consists of the fundamental spectral lines corresponding to an amplified version of the input frequency spectrum [30]. In summary, the unexpected result is achieved by adding a “bad” (less linear) amplifier to a “good” (more linear) amplifier resulting in a “better” composite amplifier.

4.2 Cancellation Conditions Using Power Series Expansion

The condition for cancellation of third order distortion in Fig. 4.1 can be derived

using power series polynomial expansions (3.1). Amplifier A1 can be modeled with a power series as shown in (4.1).

$$y_1 = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \tag{4.1}$$

Here, a_0, a_1, a_2 and a_3 describe the nonlinear behavior of amplifier A1. In the same manner, nonlinearities of amplifier A2 can be modeled with power series (4.2):

$$y_2 = b_0 + b_1 (kx) + b_2 (kx)^2 + b_3 (kx)^3 \tag{4.2}$$

Again, b_0, b_1, b_2 and b_3 describe the nonlinear behavior of amplifier A2. The input signal is x and y_1 and y_2 are the output signals of amplifiers A1 and A2, respectively. Output y_2 is subtracted from output y_1 resulting in final output y. Final output y can be described with (4.3).

$$y = y_1 - y_2 = a_0 - b_0 + a_1 x - b_1 (kx) + a_2 x^2 - b_2 (kx)^2 + a_3 x^3 - b_3 (kx)^3 . \tag{4.3}$$

As can be seen from (4.3), third order distortion is canceled and eliminated when a_3-b_3k^3=0 or $k=(a_3/b_3)^{1/3}$. In order to prevent the cancellation of the linear components (a_1-b_1k) must not equal zero. The same method can be used to eliminate other nonlinearities of order n such as second order or fifth order by setting $k=(a_n/b_n)^{1/n}$ [30].

4.3 Cancellation Condition in Terms of Output Intercept Point

It is not straight-forward to obtain coefficients of power series (4.1) and (4.2). A more practical way to express the condition of cancellation is achieved using measured output intercept points OIP3 and gains G. To derive the condition for cancellation,

consider Fig. 4.1 and let the coupling coefficient k be one, resulting in an equal distribution of the input signal between the inputs of the two amplifiers. Then, let amplifier A1 have gain G_1 in dB and third order output intercept point $OIP3_1$ in dBm, similarly amplifier A2 has a gain of G_2 and third order intercept point $OIP3_2$.

The third order output intermodulation distortion levels are denoted P_{3rd1} and P_{3rd2} for amplifier A1 and A2 respectively, and they are found by solving equation (3.28) for $P_{3rd} = 3P_{out}-2OIP3$, where P_{out} is the output power level in dBm of the linear component. P_{out} can be substituted with $P_{out}=G+P_{in}$, where P_{in} is the power level at the input of the two amplifiers. Third order intermodulation distortion levels of the two amplifiers can then be determined by equation (4.4) for amplifier A1 and equation (4.5) for amplifier A2.

$$P_{3rd1} = 3(G_1 + P_{in}) - 2OIP3_1 \tag{4.4}$$

$$P_{3rd2} = 3(G_2 + P_{in}) - 2OIP3_2 \tag{4.5}$$

In (4.4) and (4.5), $OIP3_1$ and $OIP3_2$ are the third order output intercept points of amplifier A1 and amplifier A2, respectively. G_1 is the gain of amplifier A1 and G_2 is the gain of amplifier A2. The condition for equal third order intermodulation distortion levels is reached by setting $P_{3rd1}=P_{3rd2}$ resulting in (4.6).

$$3(G_1 + P_{in}) - 2OIP3_1 = 3(G_2 + P_{in}) - 2OIP3_2 \tag{4.6}$$

(4.6) can be simplified, resulting in the simple expression (4.7) describing the cancellation condition.

$$2(OIP3_1 - OIP3_2) = 3(G_1 - G_2) \tag{4.7}$$

In (4.7), G_1-G_2 is the gain difference and must not be zero to prevent annihilation of the desired signal [31]. $OIP3_1$–$OIP3_2$ is the difference of third order output intercept points. Expression (4.7) describes the condition for cancellation of third order distortion and as one can see, (4.7) is independent of input power levels. A derivation of the theoretical gain of the composite amplifier, G_C=G_1-G_2 is shown as follows.

The subtraction at the output of the two amplifiers in Fig. 4.1 takes place at the voltage or current level. If voltages are subtracted, then the final output voltage would be v_c=v_{o1}-v_{o2}. The voltages v_{o1} and v_{o2} would be output voltages of amplifier A1 and amplifier A2, respectively, and v_c would be the output voltage of the composite amplifier. The output voltages can be expressed with the linear voltage gain g of the amplifiers, as follows

$$\begin{aligned} g_c v_{in} &= g_1 v_{in} - g_2 v_{in} \\ g_c &= g_1 - g_2 \end{aligned} \tag{4.8}$$

In above equation (4.8), g_c is the linear voltage gain of the composite amplifier, g_1 and g_2 are the linear voltage gains of amplifier A1 and amplifier A2. To find gain G_C in dBm, the logarithm with base 10 is applied to equation (4.8), resulting in expression (4.9), where G_C is the theoretical gain in dB of the composite amplifier and G_1 and G_2 are gains in dB of amplifier A1 and amplifier A2, respectively.

$$G_C = 20\log_{10}(10^{\frac{G_1}{20}} - 10^{\frac{G_2}{20}}) \tag{4.9}$$

Similarly, noise figure can be computed for the composite amplifier.

4.4 Theoretical Evaluation of Cancellation Condition

A mathematical evaluation of the condition of cancellation described with equation (4.7) is performed in Section 4.4.1 with varying input power levels using Mathcad [32]. Before proceeding to this evaluation, a basic relationship between OIP3 and 1dB compression point is established. This relationship is used in Section 4.4.1 for the mathematical evaluation with Mathcad.

Using the results from Chapter 3, a theoretical relationship can be established as shown in (4.10)

$$\frac{A_{OIP3}}{A_{1dBout}} = \frac{\sqrt{\frac{4}{3}\frac{c_1^{\,3}}{c_3}}}{0.89125 \cdot 0.3808\sqrt{\frac{c_1^{\,3}}{c_3}}} = 3.4 \tag{4.10}$$

or with expression (4.11) in terms of 1dB compression point and third order output intercept point

$$OIP3 - OP1dB = 20\log_{10} 3.4 = 10.63dB\,. \tag{4.11}$$

In equations (4.10) and (4.11), A_{OIP3} and A_{1dBout} are the amplitudes for third order intercept point and 1dB output compression point, respectively. Coefficients c_1 and c_3 are linear and third order coefficients. Furthermore, OP1dB and OIP3 are 1dB output compression point and third order output intercept point. In most practical RF systems, the difference in OP1dB and OIP3 is about 10 to 15 dB [29], typically 12 dB.

To also illustrate the theoretical value of 12 dB, a simulation of an ideal limiter was performed using Agilent Advanced Design System (ADS) [33]. The limiter has the transfer function depicted in Fig. 4.2. The input-output characteristic has an ideal linear region with voltage gain and slope of one. The output signal is limited to +1 V and –1 V. A two-tone test signal was applied to the input of the limiter and the input signal amplitudes were increased from zero to greater than the compression point. As a result, the 1dB compression point was found to be approximately 3 dBm with a third order intermodulation distortion level of –19.6 dBm. The calculated OIP3 equals 14.3 dBm, thus a difference between OP1dB and OIP3 (14.3 dBm–3 dBm) is 11.3 dB. In many circuits, OIP3 is approximately 12 dB greater than the 1dB compression point [29].

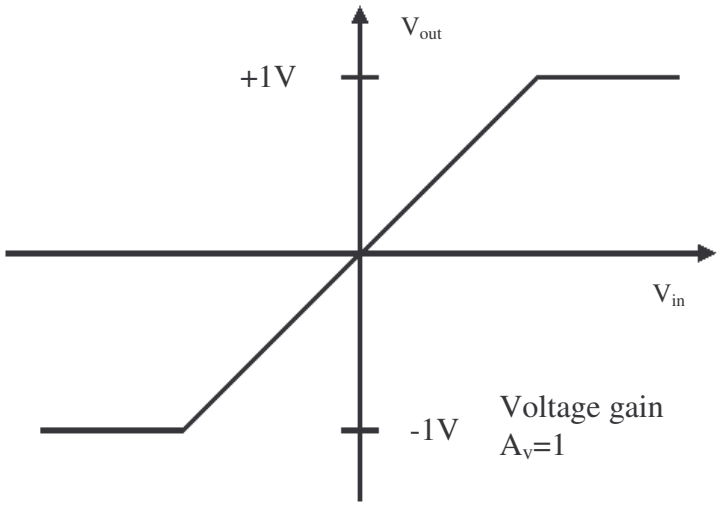

Figure 4.2 Ideal limiter voltage transfer characteristic: with linear gain of 1 and limits ± 1V at the output

4.4.1 Preliminary Mathematical Analysis of Cancellation Condition

In this section, a preliminary mathematical analysis of the cancellation condition

described with equation (4.7) is performed using Mathcad. The purpose of the evaluation is to gain insight into the performance of the new linearization technique. The Mathcad program could also be used for design purposes if the calculated results show good agreement with the measured results.

For the following calculation with Mathcad, it is assumed that the 1dB compression point of the amplifiers occurs at an output power level of (OIP3-12) dBm. Otherwise, the output power level of the amplifiers is defined as gain added to the input power in dBm, $G+P_{in}$. Fig. 4.3 illustrates the circumstances. As shown on the left side in Fig. 4.3, the input is equally distributed between more linear amplifier A1 and less linear amplifier A2. Output of amplifier A2 is subtracted from output of amplifier A1 as depicted on the right side in Fig. 4.3. Both amplifiers are characterized by parameters gain G and third order output intercept point OIP3. The subtraction of the output signals takes place at the

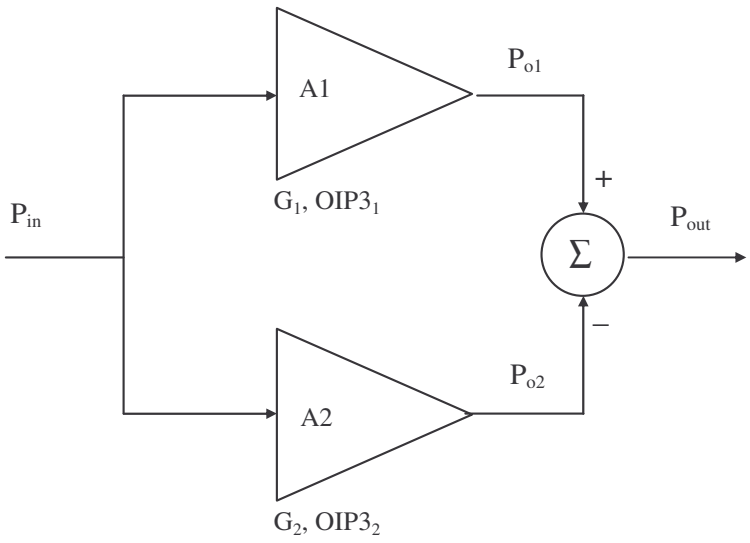

Figure 4.3 Topology of new linearization method: inputs of the two amplifiers are connected together as primary input, amplifier A1 is more linear amplifier, amplifier A2 is less linear amplifier; output signals are subtracted; gain G and third order output intercept point describe the amplifiers nonlinear characteristic [30]

voltage or current level and not the power level. The final output power P_{out} is calculated with equation (4.12)

$$P_{out} = 20\log(10^{\frac{Po1}{20}} - 10^{\frac{Po2}{20}}) . \quad (4.12)$$

The output power levels in dBm of amplifiers A1 and A2 are denoted P_{o1} and P_{o2}, respectively. The final output in dBm is P_{out}.

Independent parameters are gain, $G_{1,}$ third order intercept, $OIP3_1$ of amplifier A1 and the gain difference G_1–G_2. The gain of amplifier A2 is calculated using the gain difference and $OIP3_2$ is calculated with the use of equation (4.7). Gain and third order output intercept points of amplifier A1 and amplifier A2 from Fig. 4.3 are summarized in Table 4.1. The complete Mathcad program is attached in Appendix B.

Table 4.1 Parameter of amplifiers A1 and A2 for Mathcad calculation

	Amplifier A1	Amplifier A2
Gain	-29 dB	-40 dB
OIP3	-6 dBm	-22.5 dBm

Calculated results are discussed in the following paragraphs, using Fig. 4.4 through Fig. 4.6. First, Fig. 4.4 shows calculated linear output power level over input power level. The solid line characterizes the linear output power of the composite amplifier (A1+A2). The dashed curve describes the output power level of amplifier A1 and the dash-dotted trace depicts the output power level of amplifier A2.As reference, the third order distortion suppression is shown as dotted curve in the upper right corner of Fig. 4.4. Suppression is defined as the amount of reduction of third order intermodulation

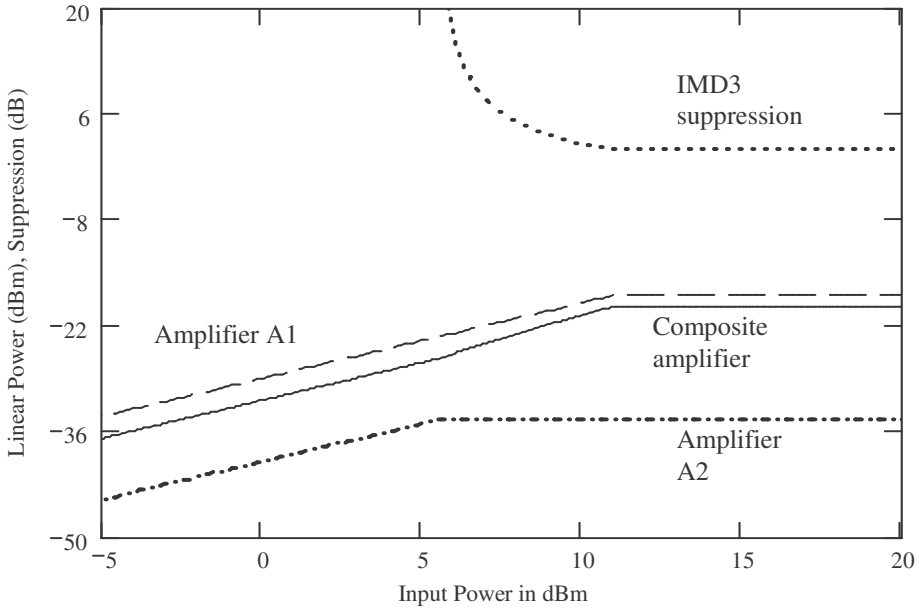

Figure 4.4 Calculated (Mathcad) output power levels: linear output power of amplifier A1 (dashed), amplifier A2 (dash-dot) and composite (solid) amplifier; suppression of third order distortion is dotted line. The input power is varied. Vertical axis displays linear output power in dBm and third order distortion suppression in dB.

product level at the final output relative to the third order intermodulation product level of amplifier A1.

A third order distortion suppression of greater than 80 dB can be reached before diminishing to a level of about 1.9 dB as shown in Fig. 4.4. Furthermore, in Fig. 4.4, third order distortion suppression begins to decrease at an input power level of 5.4 dBm. As one can see from Fig. 4.4, amplifier A2 compresses at the same input power level of 5.4 dBm whereas the output power level of amplifier A1 continues to increase up to 11 dBm input power. Due to compression of amplifier A2, the gain difference deviates from the value necessary to uphold the cancellation condition of equation (4.7). The output power level of the composite amplifier, in Fig. 4.4, shown as solid line, clips also at an input power level of 11 dBm and remains at a maximum output level of –19.3 dBm.

In addition, the final output power is reduced compared to the output power of A1. In the region between compression points of amplifier A2 and A1, the final output power increases with greater slope because less power is subtracted from the output of amplifier A1.

Secondly, to gain more insight in the mechanism of third order cancellation, third order distortion products of the two amplifiers in Fig. 4.3 and the composite amplifier are plotted in Fig. 4.5. Fig 4.5 shows the third order distortion power level versus the input power level. The solid line displays the third order distortion of amplifier A1 only, third order distortion of amplifier A2 is depicted as dashed line in Fig. 4.5 and the third order distortion at the final output is shown as dash-dotted line. As a reference, the third order distortion suppression is displayed as dotted line on the upper right corner in Fig. 4.5. As one can see in Fig. 4.5, the third order intermodulation product of amplifier A2 also compresses at an input power of 5.4 dBm whereas third order distortion of A1 continues to increase. As a result of the different third order power levels, third order products no longer cancel and third order distortion raises rapidly at the final output of the composite amplifier. This mechanism results in imperfect suppression. IMD3 of the composite amplifier assumes approximately the same value as IMD3 of amplifier A1 when A1 compresses.

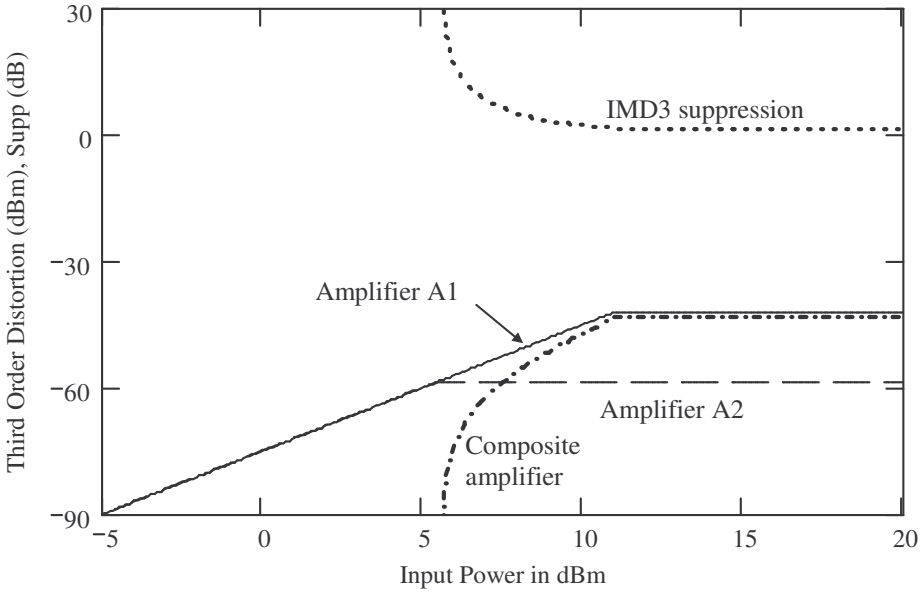

Figure 4.5 Calculated (Mathcad) third order distortion levels: third order Intermodulation distortion of A1 (solid), A2 (dashed) and composite (dash-dot) amplifier over input power; suppression of third order distortion is included as dotted line.

Gain, as one adjustable parameter in this theoretical examination, is another important parameter. Fig. 4.6 shows gain curves over the input power of composite amplifier and amplifier A1 of Fig. 4.3. The gain of amplifier A1 is described with the top dashed curve. The solid curve in Fig. 4.6 characterizes the gain of the composite amplifier.

Gain of amplifier A1 remains, as defined in the Mathcad program, constant until compression begins. The composite amplifier exhibits diminished gain due to subtraction of output power levels of the two amplifiers A1 and A2. The increase in gain of the composite amplifier is a result of subtraction of reduced output power of amplifier A2. Amplifier A1 compresses at an input power level of 11 dBm.

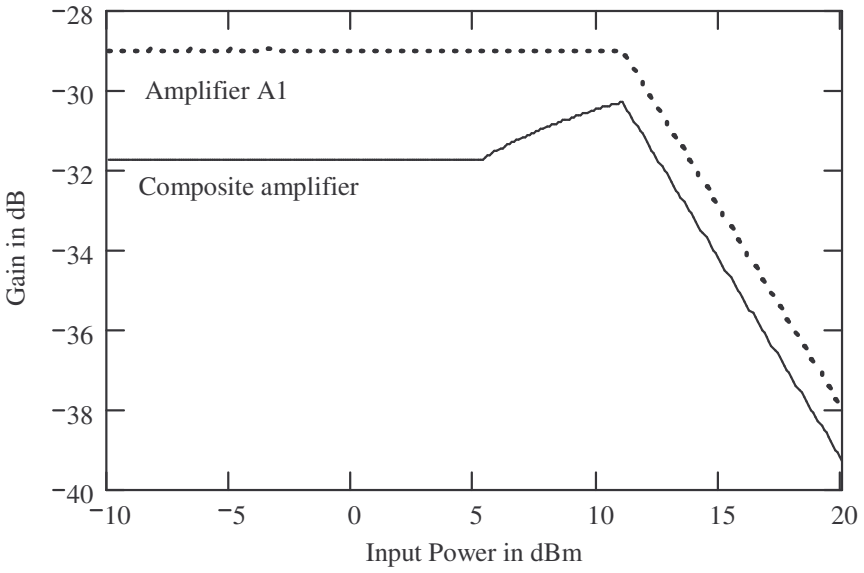

Figure 4.6 Calculated (Mathcad) gain: gain of amplifier A1 (dotted) and gain of composite amplifier (solid), G_1=-29 dB, G_C=-31.8 dB, gain of composite amplifier rises to –30.4dB.

To summarize the theoretical considerations: the new linearization technique promises an improvement in third order distortion for the topology in Fig. 4.3 for applications with low input power signals, such as receivers. Some improvement in linearity is achieved above the compression point of the less linear amplifier A2.

CHAPTER 5: INTEGRATED CIRCUIT REALIZATION OF NEW LINEARIZATION METHOD

This chapter presents data for the commonly known cross-coupled differential pair in the context of the preceding analysis. The topology of the cross-coupled differential pair is used to show the validity of equation (4.7) for integrated circuits and demonstrates the overall approach. The cross-coupled differential pair is also a basic building block for more sophisticated topologies discussed later in Chapter 7. Measured results, discussed within this chapter, confirm equation (4.7) for fully integrated circuits. This measured data forms a baseline for future research.

5.1 Cross-coupled Differential Pair

The new linearization technique lends itself to integrated circuit realization because no external parts are necessary. For a first prototype, a cross-coupled differential pair [21] was chosen as the test circuit, because of its simple subtracting mechanism. The circuit of the cross-coupled differential pair is depicted in Fig. 5.1. Compared with Fig. 4.3, the subtraction of the signals at the output of the amplifiers is accomplished by cross-coupling the drains of transistors M1 and M2 with transistors M3 and M4 in Fig. 5.1, hence the name cross-coupled differential pairs. Because drain currents are subtracted, this method of subtraction is sometimes called current differencing [21]. Current differencing is defined in [21] as the process of subtracting two nonlinear currents with equal nonlinear terms but unequal linear terms, resulting in a linear input-output

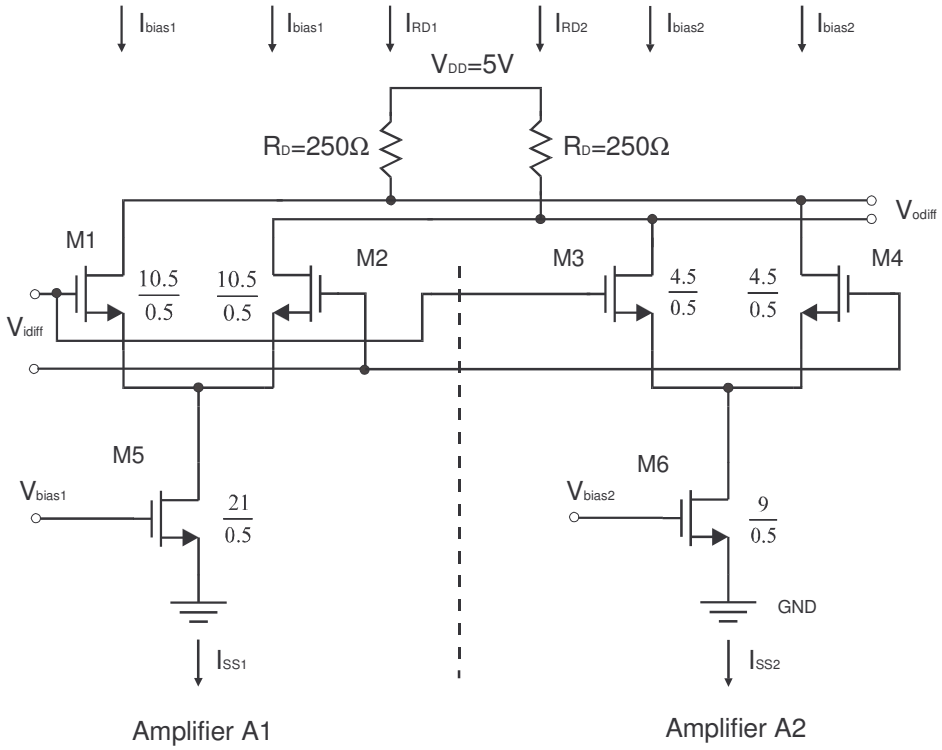

Figure 5.1 Cross-coupled differential pair: amplifier A1 consists of differential pair M1 and M2, and current source M5, amplifier A2 consists of differential pair M3 and M4 with transistor M6 as current source; drain resistors R_D=250Ω, the supply voltage is denoted V_{DD}=5V; drains of M1 and M2 are cross-coupled with drains of transistors M3 and M4. Tail currents I_{SS1} and I_{SS2} can be adjusted with V_{bias1} and V_{bias2}. Input V_{idiff} and output V_{odiff} voltages are differential voltages.

dependence. In Fig. 5.1, amplifier A1 consists of transistors M1 and M2, comprising a differential pair, and transistor M5 representing a current source. Similarly, amplifier A2 consists of a differential pair, comprised of M3 and M4, and a current source M6. All transistors are of n type channel. Both amplifiers share drain resistances R_D=250 Ω. The supply voltage equals 5 V. Drain currents I_{SS1} and I_{SS2} of M5 and M6, respectively, can be varied with bias voltages V_{bias1} and V_{bias2}, respectively. The differential input voltage is denoted V_{idiff} and the differential output voltage is denoted V_{odiff}. The input terminals

are connected parallel. The number ratio next to the transistors in Fig. 5.1 is the aspect ratio of the transistors, W/L in microns.

In order to prevent capacitive nonlinearities from interaction with the nonlinear behavior of transistors, n-diffusion resistances R_D were kept low in value. The n-diffusion resistors exhibit nonlinear capacitance due to a pn-junction between n-diffusion and p-doped substrate. As a result, the larger the n-diffusion layer is, the larger the nonlinear parasitic capacitance [34].

The small signal differential voltage gain of the cross-coupled differential pair in Fig. 5.1 is denoted A_{vdiff} and can be calculated with equation (5.1).

$$A_{vdiff} = \frac{v_{odiff}}{v_{idiff}} = -(g_{mA1} - g_{mA2})R_D \tag{5.1}$$

In (5.1), differential input voltage and output voltage are denoted v_{idiff} and v_{odiff}, respectively. Transconductance g_{mA1} is transconductance of transistors M1 and M2, and g_{mA2} is transconductance of transistors M3 and M4. R_D is the drain resistor. Equation (5.1) is valid if all transistors in Fig. 5.1 have square-law behavior. Furthermore, channel length modulation is neglected, thus drain-source resistance is infinite. A full derivation of the small signal low frequency gain is shown in Appendix C.

The large signal DC transfer characteristic of a single differential pair can be found in the literature [34] as

$$\Delta I_d = I_{d1} - I_{d2} = \sqrt{I_{SS} k_n \frac{W}{L}} V_{idff} \sqrt{1 - \frac{k_n \frac{W}{L}}{4 I_{SS}} V_{idiff}^{\ 2}}\ , \tag{5.2}$$

assuming square law behavior of the transistors comprising the differential pair. However, in the case of the presented prototype, the transistors do not exhibit square-law behavior due to velocity saturation.

The differential drain current of a single differential pair is denoted ΔI_d. If only amplifier A1 is activated, then I_{d1} and I_{d2} are drain currents of transistors M1 and M2. Length L and width W are gate length and width of transistors M1 and M2 and I_{SS} is the bias current and is denoted I_{SS1} for amplifier A1. In case only amplifier A2 is activated, then I_{d1} and I_{d2} are drain currents of transistors M3 and M4. Similarly, length L and width W are gate length and width of transistors M3 and M4 and I_{SS} is denoted I_{SS2} for amplifier A2. For equation (5.2), the differential input voltage V_{idiff}, is limited to $|V_{idiff}|=(2I_{SS}/(k_nW/L))^{0.5}$ for the transistors to remain in saturation region [34] and transconductance factor k_n equals $\mu_0 C_{ox}$ of an n-channel metal oxide semiconductor (MOS) transistor.

Equation (5.2) depicts a nonlinear function. In order to find the contributing harmonic distortion, (5.2) can be expanded into a MacLaurin series. A MacLaurin series is a Taylor series expanded around zero ($V_{idiff} = 0$). A single differential pair can be described by the following MacLaurin series [21]:

$$\Delta I_d = \sqrt{I_{SS} k_n \frac{W}{L}} V_{idiff} + 0 - \frac{1}{8} \sqrt{\frac{\left(k_n \frac{W}{L}\right)^3}{I_{SS}}} V_{idiff}^{\ 3} + 0 - \ldots \qquad (5.3)$$

The parameters and variables in (5.3) are the same as in (5.2). As one can see from equation (5.3), even order terms are zero, resulting in elimination of even order distortion, however, due to the third order coefficient, third order intermodulation

products are still present. The transfer characteristics of amplifier A1 and amplifier A2 in Fig. 5.1 can be described with equation (5.3). For proper adjustment of transistor parameters, third order distortion can also be reduced or eliminated. By setting the third order coefficient, $c_3=-0.125[(k_nW/L)^3/I_{SS}]^{0.5}$ of the two amplifiers, A1 and A2 equal, third order nonlinearities are eliminated because of subtraction of drain currents. This procedure was discussed in Chapter 4. The resulting condition for cancellation of third order nonlinearities in terms of aspect ratios and bias currents is shown in equation (5.4) [21].

$$\sqrt{\left(\frac{(W/L)_{A1}}{(W/L)_{A2}}\right)^3} = \sqrt{\frac{I_{SS1}}{I_{SS2}}} \tag{5.4}$$

In equation (5.4), $(W/L)_{A1}$ is the aspect ratio of transistors M1 and M2, and $(W/L)_{A2}$ is the aspect ratio of transistors M3 and M4. Index A1 and A2 in equation (5.4) refer to amplifier A1 and amplifier A2, respectively. The aspect ratios must not be equal to prevent cancellation of the linear term in equation (5.3). For the circuit in Fig. 5.1 the ratio can be calculated using aspect ratios of transistors M1 and M3 to $(21/9)^{3/2} = 3.56$.

In Chapter 3, a dependence of OIP3 on first and third order coefficients was derived. Equation (3.25) describes the linear output amplitude of the third order intercept point in terms of c_1 and c_3. First and third order coefficients are known for the present case of differential pairs (5.3) as shown in (5.5) and (5.6):

$$c_1 = g_m = \sqrt{I_{SS} k_n \frac{W}{L}} \tag{5.5}$$

$$|c_3| = \frac{1}{8}\sqrt{\frac{\left(k_n \frac{W}{L}\right)^3}{I_{SS}}} \tag{5.6}$$

The transconductance factor k_n is equal to $\mu_0 C_{ox}$ and W/L is the aspect ratio. I_{SS} is the bias current of the differential pair. The transconductance is denoted g_m. Substituting c_1 and c_3 in equation (3.25) with (5.5) and (5.6) leads to

$$A_{OIP3} = \sqrt{\frac{32}{3}\frac{\left(\sqrt{I_{SS} k_n \frac{W}{L}}\right)^3}{\sqrt{\frac{\left(k_n \frac{W}{L}\right)^3}{I_{SS}}}}} = \sqrt{\frac{32}{3}} I_{SS} \tag{5.7}$$

In equation (5.7), A_{OIP3} is the current level at the output third order intercept point. The important result of equation (5.7) is, that the OIP3 is only dependent on the bias current. The output power at the output third order intercept point would be calculated by multiplying $(A_{OIP3})^2$ with $R_L/2$: $P_{OIP3}=5.33 I_{SS}^2 R_L$. As one can see, the third order output intercept point is proportional to I_{SS}^2.

The integrated circuit was fabricated in a MOSIS 0.5 µm complementary metal-oxide semiconductor (CMOS) standard process and packaged in a 40-pin dual-in-line ceramic chip package [35]. Table 5.1 summarizes process parameters of the 0.5 µm CMOS process for the batch with number T2AK-AZ.

Table 5.1 Process parameters for NMOS transistor for a 0.5µm CMOS process (batch: T2AK-AZ) [35]

Threshold voltage: V_{T0}	Electron mobility: μ_0	Transconductance factor: $\mu_0 C_{ox}$	Body-Effect factor: γ	Fermi potential: Φ_F
0.75 V	475 $cm^2/V/s$	116 $\mu A/V^2$	0.48 $V^{0.5}$	0.39 V

5.2 Basic Test Setup

Before presenting the measured results, the basic test setup with the utilized measurement equipment is described. The integrated circuits' nonlinear behavior was characterized with a two-tone test. Fig. 5.2 illustrates the setup. The two sinusoidal tones are created with two signal generators of type HP E4400B, tuned to two different constant frequencies closely spaced. The signals are then combined with the help of two 10 dB attenuators in order to provide sufficient isolation between the signal generators. The combined sinusoidal tones are inputted to a third attenuator with 3 dB attenuation. The output of the attenuator is then input to the amplifier test printed circuit board (PCB). The first results presented in the following section were measured at a frequency of 110 MHz on the wafer probe station RF-1 from Cascade Microtech. For this reason, a bias tee was necessary to bias the MOS transistor at the input of the test circuit in Fig. 5.1. The bias voltage was V_{gate}=4.5 V. The pads of the chip were directly contacted with wafer probes of type ACP40 (air coplanar probe for 40 GHz bandwidth). The output signal of the amplifier under test was amplified using a Minicircuits ERA-3SM amplifier [36] mounted on a PCB, and output signal spectrum recorded with a spectrum analyzer of type HP E4401B. A capacitance between test circuit and Minicircuits amplifier makes sure

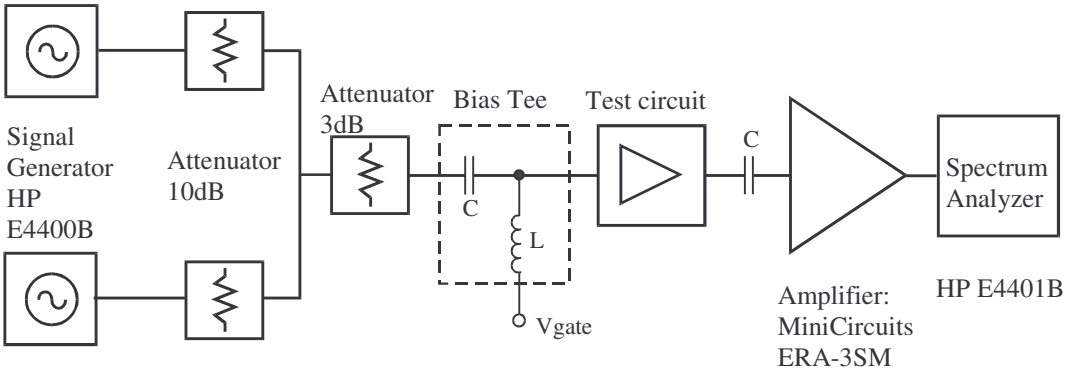

Figure 5.2 Basic test setup for two-tone measurement: on the left side, two sinusoidal signals are generated by two signal generators, the two 10dB attenuators are passive devices, the bias tee in the middle consists of a capacitance and inductance and serves as means of supplying bias voltage V_{gate} to the input of the test circuit. The test circuit is mounted on a PCB but was measured using a wafer probe station, the following device C is a coupling capacitor; the output spectrum of the test circuit is displayed with a spectrum analyzer. The input power at the test circuit is approximately 17 dB below the reading of the signal generator

that the DC operating points of the two devices are not affected. The Minicircuits amplifier with a gain of 20 dB, was added to increase the power level and overcome the restrictions of the spectrum analyzers' noise floor. All measurements were performed single-ended, with the gates of transistors M2 and M4 in Fig. 5.1 biased with 4.5 V, but connected to small-signal ground via a 100 nF capacitor.

5.3 Measured Results

In the following section, measured results are presented and discussed. In order to prove the theory of the new linearization method, in particular equation (4.7), third order intermodulation product and output power levels of the output signal were measured and recorded. Third order output intercept points were calculated from the measured values. The third order cancellation performance was measured at frequencies 1, 3, 10, 30 and 110 MHz. The third order cancellation performance at higher frequencies was reduced,

due to bandwidth limitations of chip package and signal leakage between input and output of the test circuit. The displayed power level on the signal generators was taken as input power level, which is with respect to a matching 50 Ω load. Due to attenuation, the input power at the test circuit is approximately 17 dB below the reading of the signal generator. During the measurement of amplifier A1, amplifier A2 was turned off and during the measurement of amplifier A2, amplifier A1 was turned off. For the composite amplifier, both amplifiers A1 and A2 were turned on and adjusted for cancellation of third order intermodulation products.

5.3.1 Measured Results for 110 MHz Tested with Wafer Probe Station

First, output frequency spectra for a test frequency of 110 MHz are shown in Fig. 5.3, 5.4, 5.5 and 5.6. This set of plots depicts frequency spectra of more linear amplifier A1, less linear amplifier A2, composite amplifier, and signal leakage. Signal leakage was measured between input and output when both amplifiers were deactivated. The input power level for the set of figures for the present case was 0 dBm at the signal generator and –17 dBm at the input of the test circuit. Similarly, the reading at the spectrum analyzer is approximately 20 dB higher than the output power of the test circuit.

In Fig. 5.3, the two spectral lines in the center are the fundamental frequencies at 110 MHz and 110.1 MHz and with output power levels of –29.2 dBm. Third order intermodulation products of –75.7 dBm power level appear at frequencies 109.9 MHz and 110.2 MHz, depicted as outermost spectral lines in Fig. 5.3. Fig. 5.4 pictures the output frequency spectrum of amplifier A2. Fundamental tones have a power level of –40.6 dBm and third order distortion products have a power level of –75.9 dBm. Due to proper adjustment of drain currents of the two amplifiers, A1 and A2, third order

distortion power levels are almost equal. By subtracting the two output signals, reduction in third order distortion was achieved. Fig. 5.5 shows the result in the output spectrum of the composite amplifier. In Fig. 5.5, the third order intermodulation products were reduced below the noise floor of the spectrum analyzer to approximately –98.4 dBm. Comparing third order distortion of Fig. 5.3 and Fig. 5.5, a third order suppression of approximately 23 dB was attained. The fundamental spectral lines, in Fig. 5.5 have a measured power level of –31.2 dBm. Again, comparing Fig. 5.5 with the spectrum of amplifier A1 in Fig. 5.3, a loss in gain of approximately 2 dB was measured. This result is an effect of subtracting the two fundamental signals of Fig. 5.3 and Fig. 5.4.

The signal leakage between input and output of the test circuit is shown in Fig. 5.6. If both amplifiers are turned off, a power level at the fundamental frequencies of approximately -56 dBm is still available at the output. The leakage is most likely caused by parasitic capacitances in the pad frame of the integrated circuit. Parasitic capacitances can appear between signal line and power and ground rings in the pad frame. The leakage is significantly below the output power level of the less linear amplifier A2.

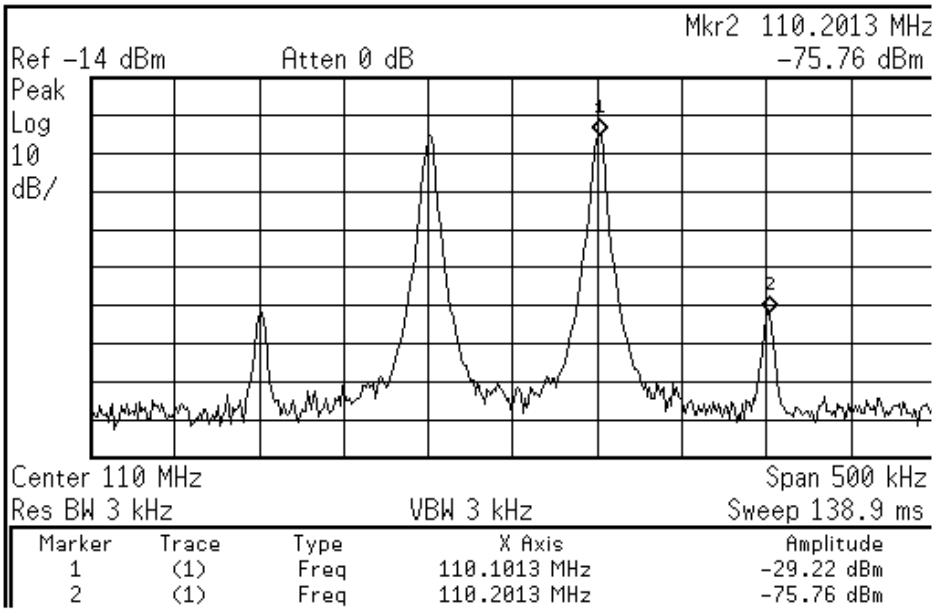

Figure 5.3 Measured output frequency spectrum of amplifier A1 at 110MHz: amplifier A2 is turned off. Fundamental signals at 110MHz and 110.1MHz and third order distortion at frequencies 109.9MHz and 110.2MHz. Output power is measured at output of Fig. 5.2, true output power is 20dB lower, circuit of Fig. 5.1 at current 0.287mA.

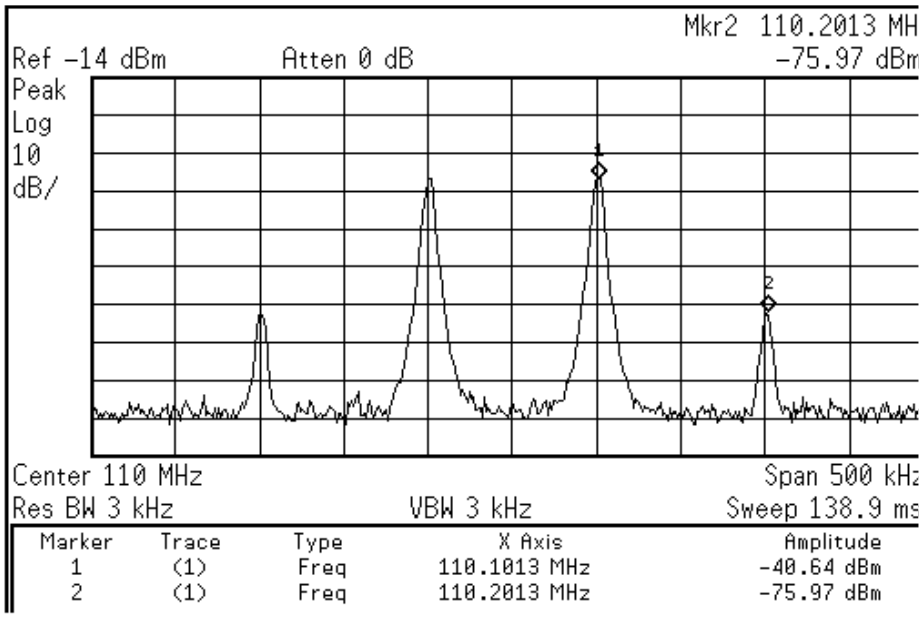

Figure 5.4 Measured output frequency spectrum of amplifier A2 at 110MHz: amplifier A1 is turned off. Fundamental signals at 110MHz and 110.1MHz and third order distortion at 109.9MHz and 110.2MHz. Output power is measured at output of Fig. 5.2, true output power is 20dB lower, circuit of Fig. 5.1 at current 0.026 mA.

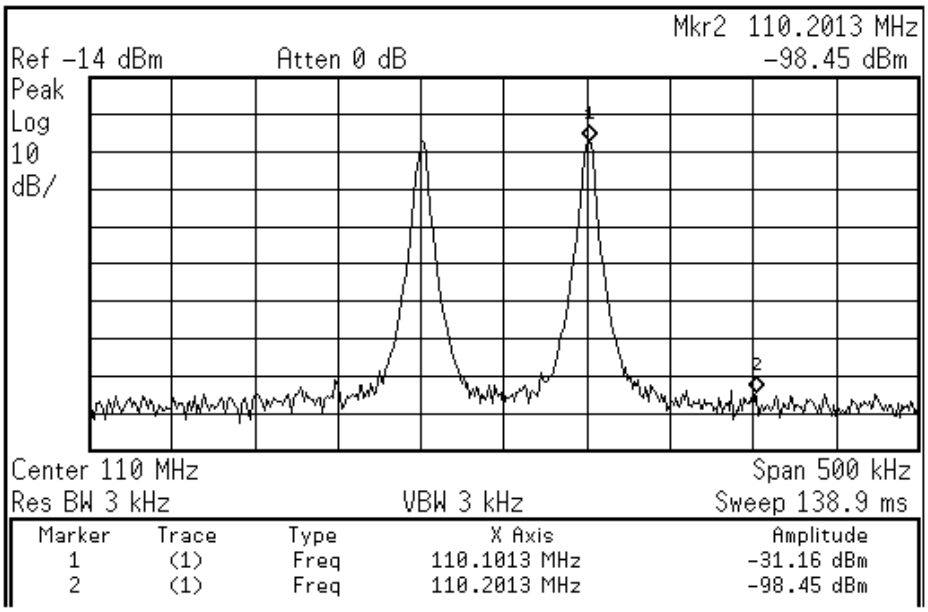

Figure 5.5 Measured output spectrum of composite amplifier at 110MHz showing cancellation of third order distortion at frequencies 109.9MHz and 110.2MHz. Output power is measured at output of Fig. 5.2, true output power is 20dB lower, circuit of Fig. 5.1 at currents 0.286 and 0.026 mA

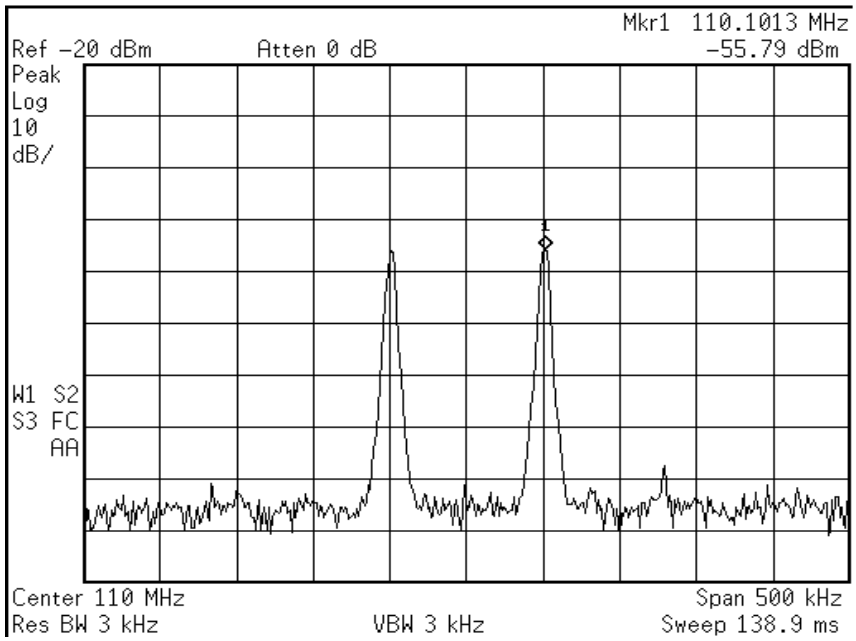

Figure 5.6 Measured signal leakage between input and output of test circuit at 110MHz: both amplifiers are turned off. Fundamental frequencies at 110 and 110.1MHz; Results measured using test setup of Fig. 5.2.

Table 5.2 summarizes the results for the set of frequency spectra for a test frequency of 110 MHz using the test setup from Fig. 5.2. The first column lists the input power P_{in}, displaying the reading of the signal generator. The input power at the test circuit is 17 dB lower. The fundamental output power, denoted P_{out}, measured at the

Table 5.2 Summary of measured results for 110 MHz using setup of Fig. 5.2

	P_{in} (dBm) (at generator)	P_{out} (dBm) (at spectrum analyzer)	G (dB)	P_{3rd} (dBm)	OIP3 (dBm)
Amplifier A1 (Fig. 5.3)	0	-29.2	-29.2	-75.8	-6.0
Amplifier A2 (Fig. 5.4)	0	-40.6	-40.6	-76.0	-23.0
Composite amplifier A1+A2 (Fig. 5.5)	0	-31.2	-31.2	-98.4	2.5
Comparison: A1 with composite amp			Gain loss: 2 dB	Third order distortion suppression: 22.6 dB	Improvement of OIP3: 8.5 dB

spectrum analyzer is shown in the second column. The linear power at the output of the test circuit is approximately 20 dB lower. Column number three from the left, shows the gain, denoted G. The gain is calculated as P_{out}-P_{in} using measured value from Table 5.2. Furthermore, third order distortion level, denoted P_{3rd}, as displayed at the spectrum analyzer is collected in column four from the right side. The third order output intercept point OIP3 is calculated with equation (3.28) using measured results and is displayed in the column on the right in Table 5.2. Furthermore, the first row on the top shows

measured results of amplifier A1 collected from Fig. 5.3, the second row from the top in Table 5.2 lists results of amplifier A2 from Fig. 5.4. Finally, the third row lists results of the composite amplifier taken from Fig. 5.5. The row at the bottom of Table 5.2 shows comparisons of the results of amplifier A1 and composite amplifier.

The measured results show an improvement in third order distortion of 22.6 dB and an improvement of OIP3 of 8.5 dB. A gain loss of 2 dB of the composite amplifier relative to amplifier A1 was measured. In addition, the measured results in Table 5.2 can be used to prove expression (4.7) which is stated here again for review: $2(OIP3_1-OIP3_2)=3(G_1-G_2)$. Substituting the measured results into the left side of equation (4.7) leads to:

$$2(OIP3_1-OIP3_2) =2(-6.0+23.0) = 34.0dB$$

and substituting results into the right side of equation (4.7) gives:

$$3(G_1-G_2)= 3(-29.2+40.6) = 34.2dB$$

The results of both sides are almost equal, showing proof of the validity of equation (4.7).

The bias conditions of the test circuit in Fig. 5.1 for cancellation of third order distortion are summarized in Table 5.3. A combination of tail currents I_{SS1}=287 μA and I_{SS2}=26 μA, with the proper aspect ratios, leads to reduction of third order distortion which can be proven by employing equation (5.4): $(W/L)_{A1}$=21, $(W/L)_{A2}$=9, hence $(21/9)^{3/2}$=3.56 and $(I_{SS1}/I_{SS2})^{0.5}=(287/26)^{0.5}$=3.32. The discrepancy between the two values 3.56 and 3.32 could be explained with the low bias voltage of 0.867 V of transistor M6. Transistor M6 is most likely at the low end of strong inversion approaching moderate inversion, thus no longer follows the transistor square-law characteristic.

Table 5.3 Summary of bias conditions for cancellation of third order distortion

	V_{bias}	I_{SS}
Amplifier A1 (M5 in Fig. 5.1)	1.06 V	287 μA
Amplifier A2 (M6 in Fig. 5.1)	0.867 V	26 μA
Composite amplifier (M5+M6 in Fig. 5.1)	Total bias-current: 303 μA	
Bias voltage V_{Gate} of M1 through M4	4.5 V	

In Chapter 4, theoretical considerations indicate that suppression of third order distortion is reduced at higher input power levels. In order to verify a decrease in third order suppression, the input power level was varied from –10 dBm to +10 dBm. Again, all three amplifiers, A1, A2 and composite amplifier, were tested. The test frequency was maintained at 110 MHz. Fig. 5.7 depicts linear output power levels and third order intermodulation products versus the input power level taken from the signal generator. The input power at the test circuit is 17 dB lower. The output power levels are recorded at the spectrum analyzer and are 20 dB higher than the output power of the test circuit. The top three curves in Fig. 5.7 describe the characteristic of the linear output power. The output power level of amplifier A1 is denoted P_{out1}; output power of amplifier A2 is denoted P_{out2} and output power of the composite amplifier is denoted P_{outC}. The linear output of composite amplifier, P_{outC}, is approximately 2 dB lower than P_{out1} up to approximately 1 dBm input power. As previously discussed, the loss in output power is a result of subtraction of the linear fundamental components.

Due to compression of A2, beginning at an input power level of 1 dBm, less signal is subtracted from the more linear amplifiers' (A1) output signal and the final output power of the test circuit, P_{outC}, increases faster and approaches power level P_{out1}. The 1dB input compression points of amplifier A1 and A2 were measured and found to be 9 dBm and 4 dBm, respectively, using test setup in Fig. 5.2. The 1dB compression points were taken from the signal generator display. The 1dB compression point of the composite amplifier is approximately equal to that of amplifier A1.

The remaining three curves, at the bottom of Fig. 5.7, illustrate third order intermodulation distortion, IMD3, of A1, A2 and composite amplifier. For an increase in input power of 1 dBm, IMD3 increases by 3 dBm for A1 and A2. Good cancellation of third order distortion can be achieved for low input power levels up to about 3 dBm. As can be seen from Fig. 5.7, as output power of A2 compresses, third order distortion cancellation is diminished because third order intermodulation product of A2 also compresses. As a result, third order distortion levels are no longer equal for higher input power levels. Suppression is almost zero as amplifier A1 reaches its 1dB input compression point of 9 dBm, however, there is still an improvement in suppression at the 1dB compression point of amplifier A2 of about 18 dB. These measured results agree with the calculated behavior.

As an important measure for nonlinearity, the characteristic of the third order output intercept point, OIP3, over input power is shown in Fig. 5.8. The upper trace is the third order output intercept point of the composite amplifier depicted as solid line in Fig. 5.8. The dotted curve displays the characteristic of the third order output intercept point of amplifier A1 and the dashed curve at the bottom of Fig. 5.8 shows OIP3 of amplifier

A2. Again, the input power in Fig. 5.8 is taken from the signal generator and is 17 dB lower at the test circuit input. As a result of suppression of third order distortion, OIP3 is improved by more than 12 dB for low input power levels. An improvement of about 7 dB is measured at the 1dB compression point of A2 (input referred 4 dBm).

Information about the behavior of the gain of the test circuit is given in Fig. 5.9. Fig. 5.9 shows the gain of amplifiers A1, A2 and composite amplifier versus the input power. In Fig. 5.9, the gain is calculated using input power levels taken from the display of the signal generator and the output power recorded at the spectrum analyzer. The upper dotted curve in Fig. 5.9 displays the gain of amplifier A1, the solid trace describes the gain behavior of the composite amplifier. The gain characteristic of amplifier A2 is shown at the bottom of Fig. 5.9 as dashed curve. As expected, the gain of the composite amplifier is lowered by 2 dB because of subtraction of the two fundamental output signals of A1 and A2. As previously explained, the composite gain increases as A2 compresses due to the fact that less output signal of A2 is subtracted from output signal of A1. In Fig. 5.9, the overall gain finally decreases as amplifier A1 compresses.

To summarize, the linearization technique implemented in this topology cancels third order distortion for input power levels below the 1dB compression point of the less linear amplifier. Some improvement in third order distortion is achieved at input power levels between 1dB compression points of more linear and less linear amplifiers.

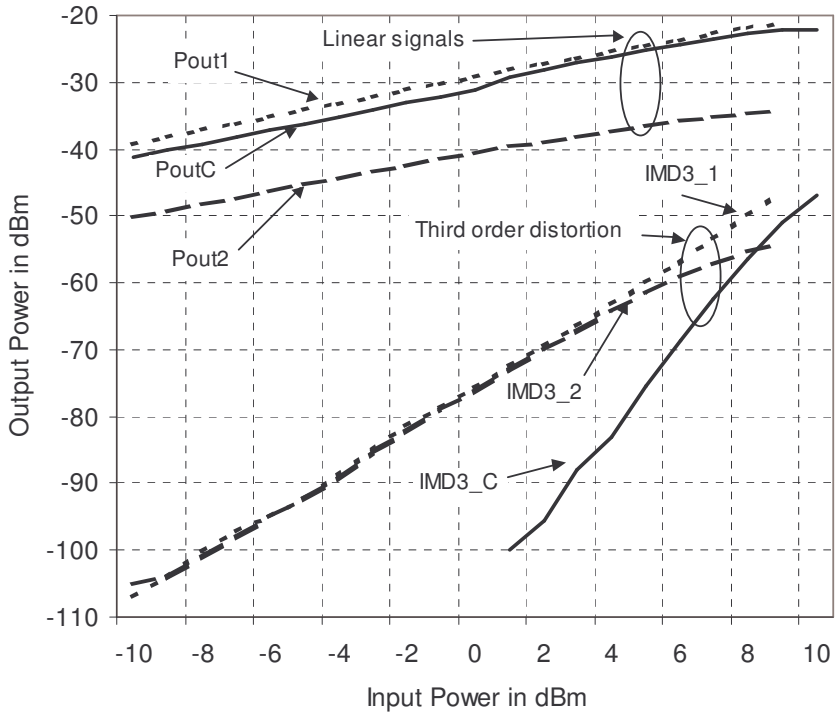

Figure 5.7 Measured linear signals and third order products of A1, A2 and composite amplifier for 110MHz over input power: Pout1 is output power of A1, Pout2 is output power of A2 and PoutC is output power of composite amplifier; third order distortion is denoted IMD3, index 1, 2 and C refer to amplifier A1, amplifier A2 and composite amplifier. Output power is measured at output of Fig. 5.2, true output power is 20 dB lower. Input power is at signal generators of Fig. 5.2, true input power is 17 dB less, circuit of Fig. 5.1 at currents 0.287 and 0.026 mA.

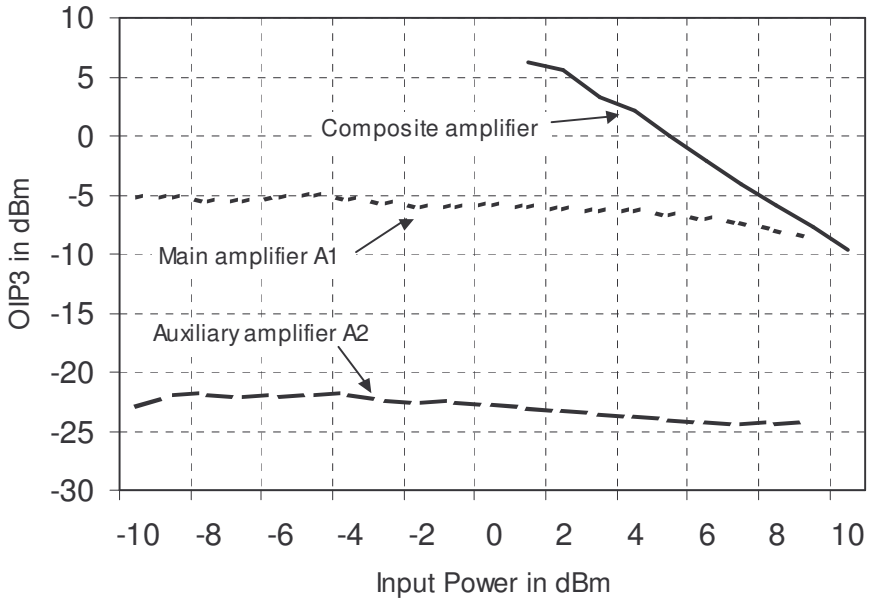

Figure 5.8 Measured third order output intercept point OIP3 of amplifier A1, A2 and composite amplifier at 110MHz over input power. Input power is at signal generators of Fig. 5.2, true input power is 17 dB less.

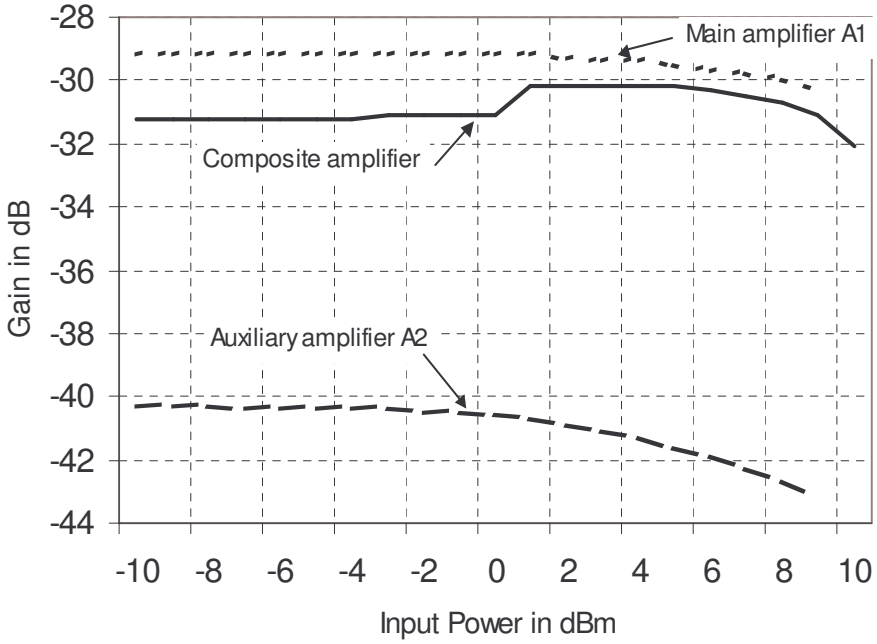

Figure 5.9 Measured gain of A1, A2 and composite amplifier at 110MHz over input power; input 1dB compression points, $P1dB_1$=9dBm, $P1dB_2$=3dBm, $P1dB_C$=9dBm. Output power is at output of Fig. 5.2, true output power is 20 dB lower. Input power is at signal generators of Fig. 5.2, true input power is 17 dB less.

5.3.2 Measured Results for Additional Frequencies

Given the foregoing behavior of the performance of cancellation of third order distortion with various input power levels, the question arises how the new method of cancellation performs over a range of frequency. To answer this question, test results for different frequencies, including 1, 3, 10 and 30 MHz are discussed below.

The following measurement results are based on a slightly different test setup from Fig. 5.2. The measurements were performed with the test circuit on a PCB; no wafer probe station was used. Fig. 5.10 illustrates the new test setup. A two-tone test was employed in the test circuit. On the left side in Fig. 5.10, two test signal generators generate two sinusoidal tones. Both tones are combined using two 10 dB attenuators to achieve good isolation between the generators. The test signal, after passing a 3 dB attenuator, is input to the ac-coupled test circuit. The RF output signal is also ac-coupled

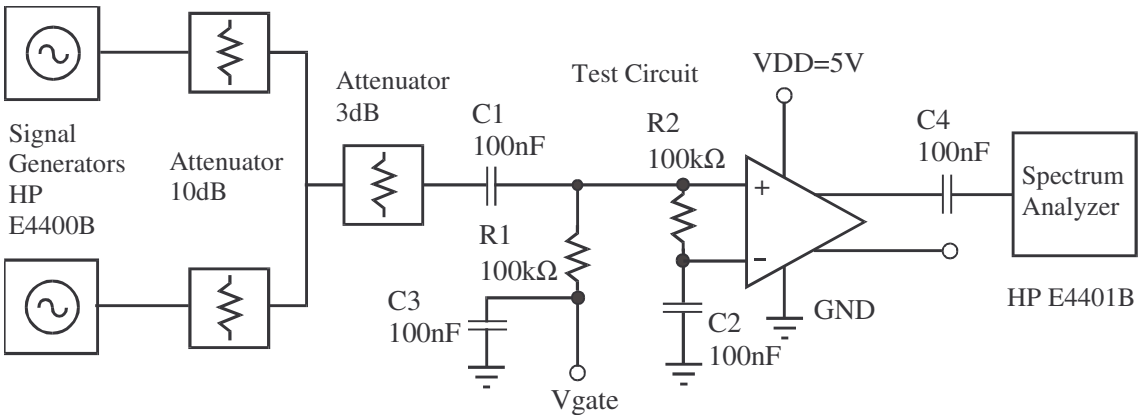

Figure 5.10 Modified test setup for two-tone measurements: on the left side, two signal generators provide two sinusoidal signals, then the two signals are combined using two passive 10dB attenuators, the following device is a 3dB passive attenuator; the input signal is ac-coupled by a capacitor C1 and input to the test circuit. The test circuit is mounted on a PCB. Resistors R1 and R2 are used for biasing the test circuit; the output signal is ac-coupled and input to a spectrum analyzer of type HP E4401B.

and input to a spectrum analyzer. The differential input of the test circuit is biased with voltage V_{gate}=4.5 V. The negative input is set to ac-ground by utilizing a capacitor C2=100 nF. Resistors R1 and R2 in Fig. 5.10 provide isolation for the RF signal. Measurements are performed single-ended. The input power level at the test circuit was kept constant at –10.1 dBm for all following presented measured results. The bias currents for the composite amplifier for the present set and for all following sets of frequency spectra for different frequencies were adjusted to I_{SS1}=300 μA for amplifier A1 and I_{SS2}=30 μA for amplifier A2. The total DC current equals 330 μA.

First, a set of frequency spectra with test results for a frequency of 1 MHz is presented similar to the set of frequency spectra for a test frequency of 110 MHz. Fig. 5.11 shows a plot of the output signal of amplifier A1. Fundamental signals appear at 1 MHz and 1.05 MHz with a power level of –45.1 dBm, shown as the two spectral lines in the center of Fig. 5.11. Third order intermodulation products appear at frequencies 0.95 MHz and 1.1 MHz with a power level of –85 dBm. Another plot for the less linear amplifier, A2, is given in Fig. 5.12. The output power level of the fundamental tones is –61.2 dBm and third order distortion is –84.6 dBm. After subtraction of the output signals, the output signal spectrum of the composite amplifier results in a spectrum, shown in Fig. 5.13. The third order distortion is reduced below the noise level of the spectrum analyzer and a third order suppression of at least 25 dB was achieved. The fundamental output power was diminished by about 1.5 dB. The input power level at the test circuit was kept constant at –10.1 dBm. Finally, Fig. 5.14 shows the signal leakage of the fundamentals with a value of -85 dBm. For the measurement of the leakage, both amplifiers were

turned off. As expected, the leakage at 1 MHz is significantly lower than the leakage at a frequency of 110 MHz, -55.8 dBm.

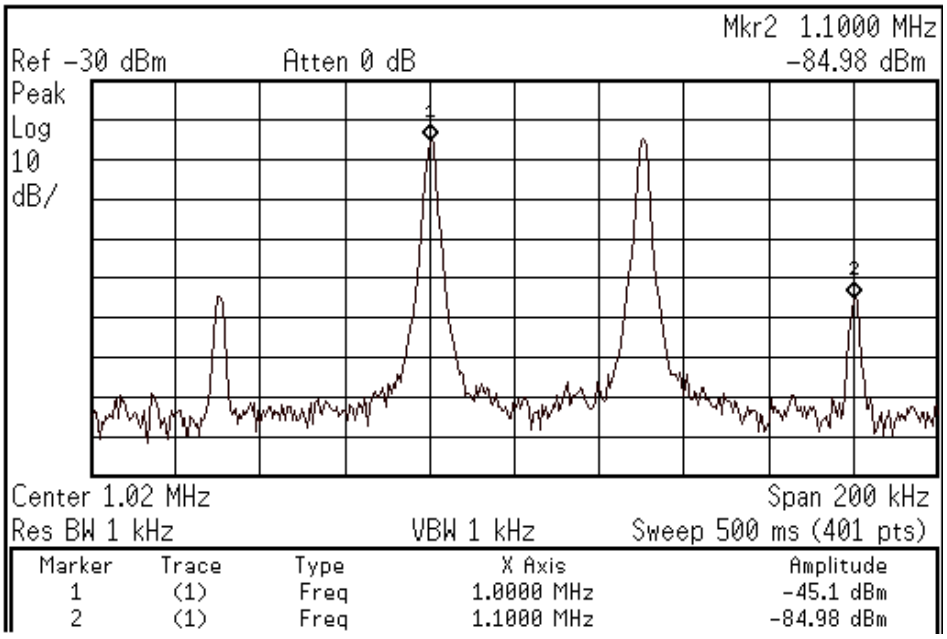

Figure 5.11 Measured output frequency spectrum of amplifier A1 at 1MHz: amplifier A2 was turned off; fundamental tones at 1MHz and 1.05MHz, third order distortion at 0.95MHz and 1.1MHz. Results measured using test setup of Fig. 5.10.

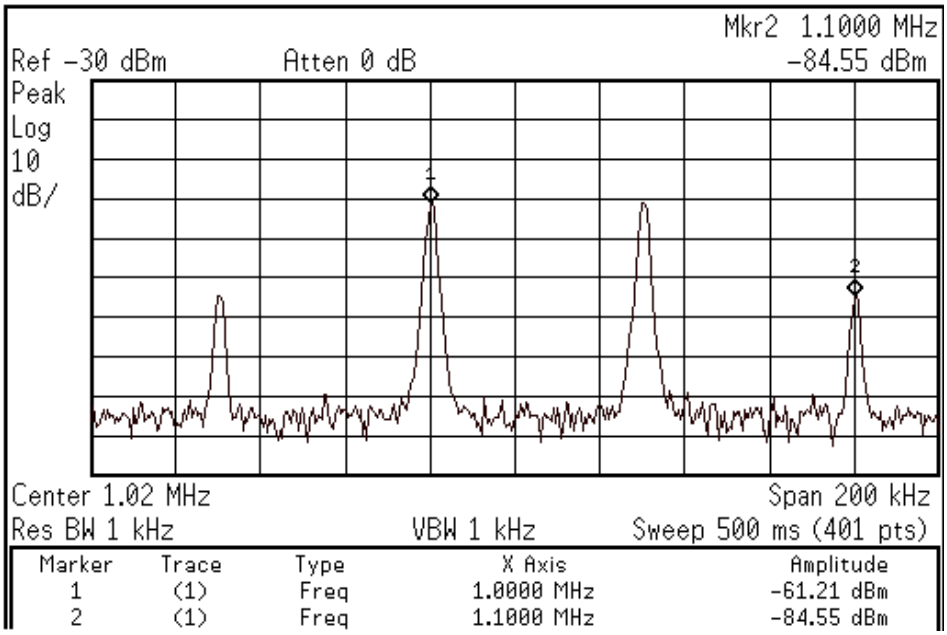

Figure 5.12 Measured output frequency spectrum of amplifier A2 at 1MHz: fundamental tones at 1MHz and 1.05MHz, third order distortion at frequencies 0.95MHz and 1.1MHz. Results measured using test setup of Fig. 5.10.

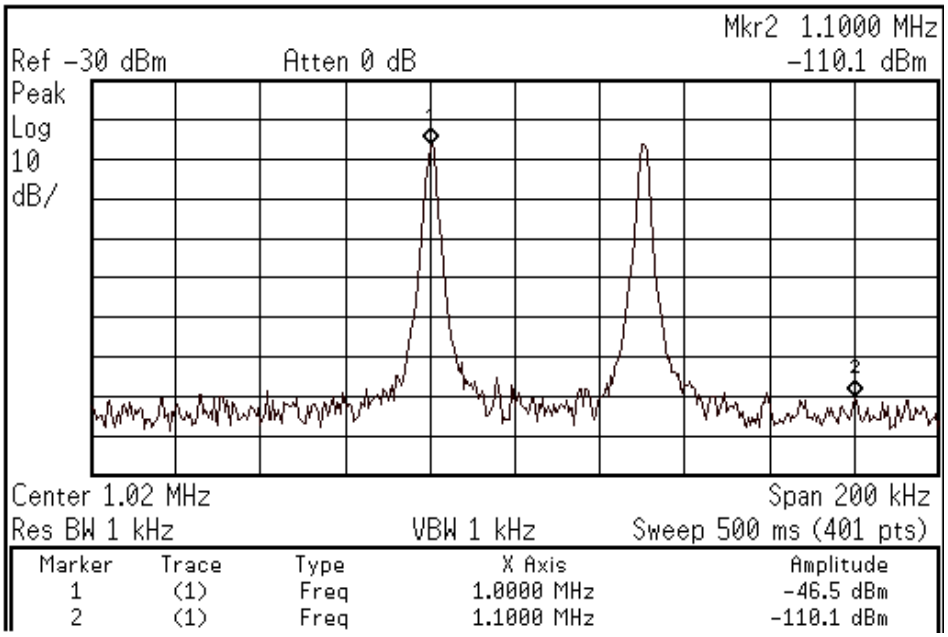

Figure 5.13 Measured output frequency spectrum of composite amplifier at 1MHz showing cancellation of third order distortion. Results measured using test setup of Fig. 5.10.

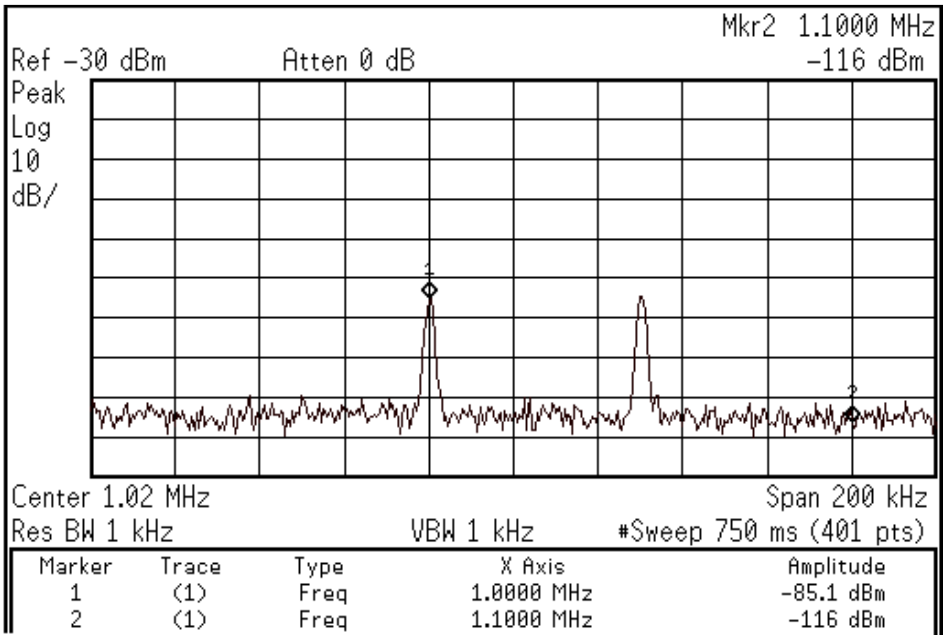

Figure 5.14 Measured signal leakage between input and output of test circuit at 1MHz: fundamental tones appear at 1MHz and 1.05MHz with power level of –85.1dBm. Results measured using test setup of Fig. 5.10.

Table 5.4 Summary of measured results for 1MHz using setup from Fig. 5.10

	P_{in} (dBm)	P_{out} (dBm)	G (dB)	P_{3rd} (dBm)	OIP3 (dBm)
Amplifier A1 (Fig. 5.11)	-10.1	-45.1	-35.0	-85.0	-25.2
Amplifier A2 (Fig. 5.12)	-10.1	-61.2	-51.1	-84.6	-49.5
Composite amplifier A1+A2 (Fig. 5.13)	-10.1	-46.5	-36.4	-110.1	-14.7
Comparison: A1 with comp. amp.			Gain loss: 1.4 dB	Third order suppression: 25.1 dB	Improvement of OIP3: 10.9 dB

Table 5.4 summarizes the measured results for a frequency of 1 MHz similar to Table 5.2.For a frequency of 1 MHz, third order distortion is suppressed by 25 dB resulting in an improvement of approximately 11 dB in OIP3 as shown in Table 5.4. Substituting the measured result from Table 5.4 into equation (4.7), similar to the case of a test frequency of 110 MHz leads to the following calculations:

2(OIP31-OIP32)=3(G1-G2)

2(OIP31-OIP32)=2(-25.2+49.5)=48.6dB

3(G1-G2)=3(-35.0+51.1)=48.3dB

The results show that expression (4.7) is also valid for a test frequency of 1 MHz.

Next, a set of measured results for a frequency of 3 MHz is presented. The set also consists of 4 plots with frequency spectra of output signals of A1, A2, composite amplifier and signal leakage. As before, a two-tone test is applied where each tone has a power level of –10.1 dBm at the input of the test circuit. First, Fig. 5.15 pictures the

frequency spectrum of the output signal of amplifier A1 when A2 is turned off. The fundamental tones at 3 MHz and 3.05 MHz have a power level of –44.8 dBm, whereas the third order distortion products have a power level of –84.2 dBm at frequencies 2.95 MHz and 3.1 MHz. Secondly, the output frequency spectrum of amplifier A2, when A1 is turned off, is shown in Fig. 5.16. Fundamental linear components have power levels of about –60.7 dBm and third order distortion levels are –84.1 dBm. The third frequency spectrum, in this present set of frequency spectra, Fig. 5.17, depicts the output frequency spectrum of the composite amplifier adjusted for cancellation of third order distortion. The fundamental components show a power level of –46.2 dBm. Compared to Fig. 5.15, fundamental power levels are reduced by about 1.4 dB. Third order components are not visible anymore; their power level was reduced below the noise floor of the spectrum analyzer. An improvement in third order distortion of at least 29 dB was achieved. The signal leakage for 3 MHz is shown in Fig. 5.18. With both amplifiers turned off, a signal leakage of the fundamental tones of –75.6 dBm is still present. The measured results are collected in Table 5.5 similar to Table 5.2.

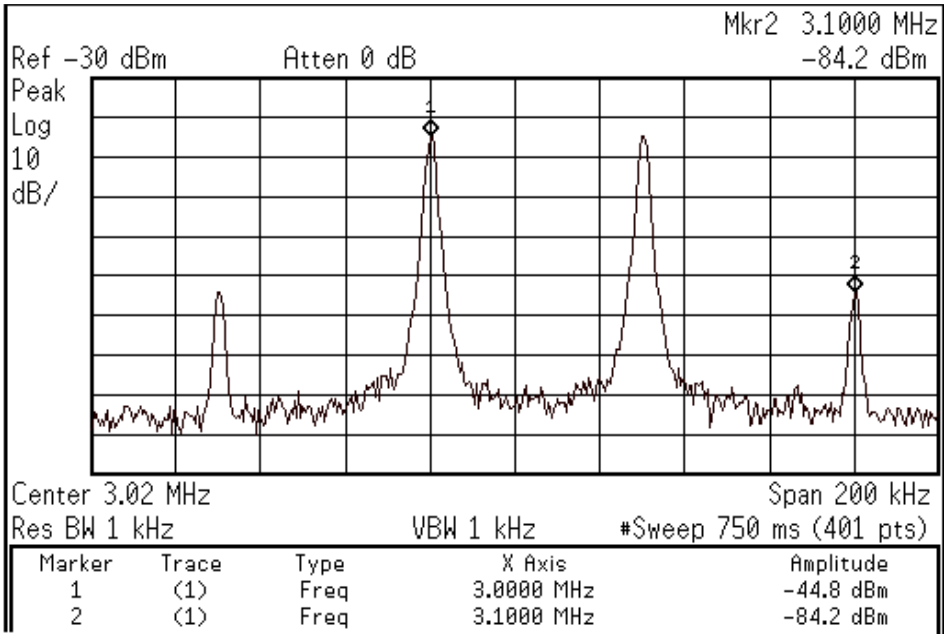

Figure 5.15 Measured output frequency spectrum of amplifier A1 at 3MHz: amplifier A2 is turned off; fundamental tones at 3MHz and 3.05MHz, third order distortion at frequencies 2.95MHz and 3.1MHz. Results measured using test setup of Fig. 5.10.

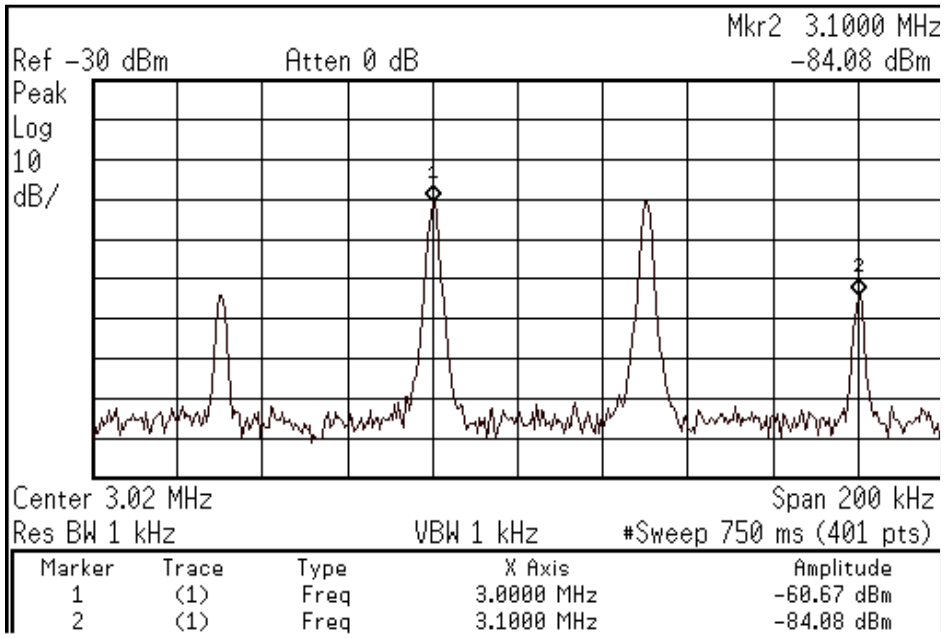

Figure 5.16 Measured output frequency spectrum of amplifier A2 at 3MHz: amplifier A1 is turned off; fundamental tones at fundamental tones at 3MHz and 3.05MHz, third order distortion at frequencies 2.95MHz and 3.1MHz. Results measured using test setup of Fig. 5.10.

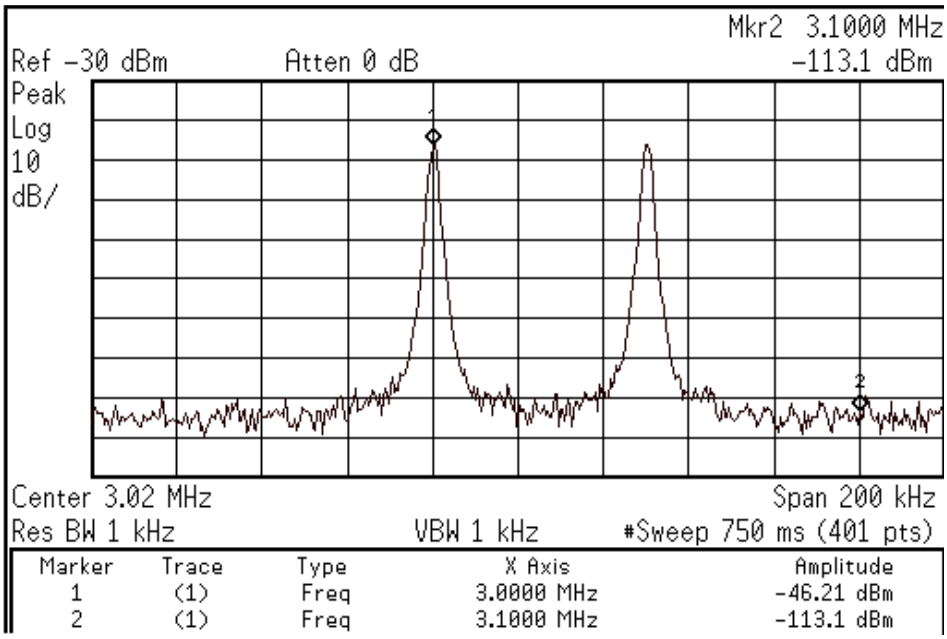

Figure 5.17 Measured output frequency spectrum of composite amplifier at 3MHz showing cancellation of third order distortion. Results measured using test setup of Fig. 5.10.

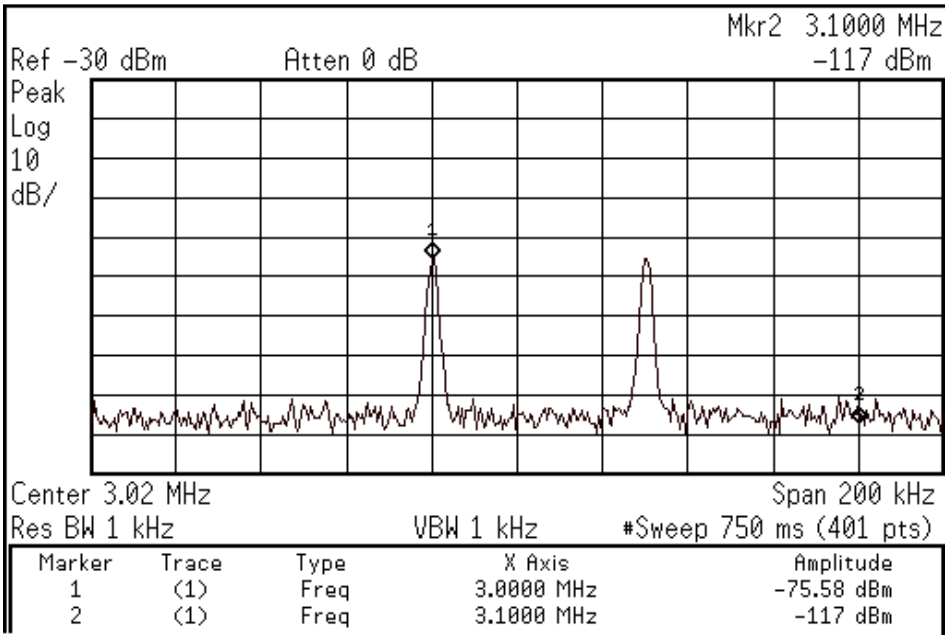

Figure 5.18 Measured signal leakage between input and output of test circuit at 3MHz; fundamental tones at frequencies 3MHz and 3.05MHz. Results measured using test setup of Fig. 5.10.

Table 5.5 Summary of measured results for 3 MHz using setup from Fig. 5.10

	P_{in} (dBm)	P_{out} (dBm)	G (dB)	P_{3rd} (dBm)	OIP3 (dBm)
Amplifier A1 (Fig. 5.15)	-10.1	-44.8	-34.7	-84.2	-25.1
Amplifier A2 (Fig. 5.16)	-10.1	-60.7	-50.6	-84.1	-49.0
Composite amplifier A1+A2 (Fig. 5.17)	-10.1	-46.2	-36.1	-113.1	-12.8
Comparison: A1 with comp. Amp.			Gain loss: 1.4 dB	Third order suppression: 28.9 dB	Improvement of OIP3: 12.3 dB

The measured results of Table 5.5 show a reduction of third order distortion of approximately 29 dB leading to an improvement of 12.3 dB in OIP3. Furthermore, a loss of 1.4 dB in gain is measured. Furthermore, equation (4.7) is valid for the integrated circuit tested with a frequency of 3 MHz as shown in the following calculations using measured results from Table 5.5:

$2(OIP3_1-OIP3_2)= 3(G_1-G_2)$

$2(OIP3_1-OIP3_2)= 2(-25.1+49.0)=47.8dB$

$3(G_1-G_2)=3(-34.7+50.6)=47.7dB$

For the third set of frequency spectra, a two-tone test with frequencies 10 MHz and 10.05 MHz and input power level of –10.1 dBm was deployed. The output frequency spectrum of amplifier A1 is presented first in Fig. 5.19. In Fig. 5.19, –44.2 dBm power levels were measured for the fundamental signals, and –81.8 dBm was measured for third order distortion products. Third order distortion products appear at 9.95 MHz and 10.1

MHz. Next, the frequency spectrum at the output of amplifier A2 is illustrated in Fig. 5.20. Fundamental tones show a power level of –59.1 dBm and third order distortion levels have a value of –82.2 dBm. Third order distortion levels of A2 are almost equal to third order distortion levels of amplifier A1. Activating both amplifiers A1 and A2 results in the composite amplifier with output frequency spectrum depicted in Fig. 5.21. Again, third order distortion is reduced by about 30 dB. Fundamental tones have a power level of –45.6 dBm. The overall gain suffers a reduction of approximately 1.4 dB. Finally, the amount of signal leakage was measured as –64.7 dBm and is shown in Fig. 5.22. Third order output intercept point and gain are summarized together with measured results in Table 5.6.

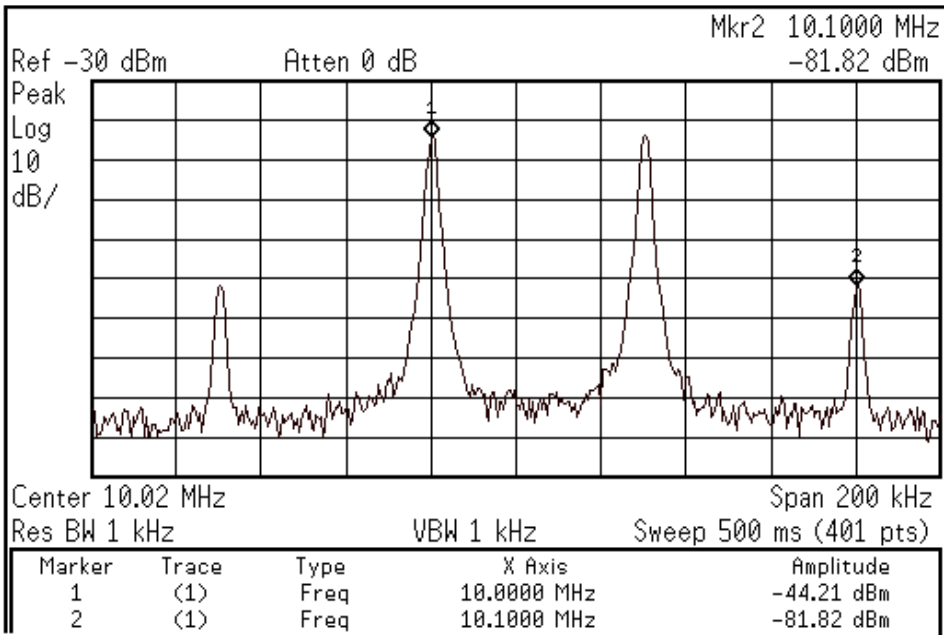

Figure 5.19 Measured output frequency spectrum of amplifier A1 at 10MHz: amplifier A2 is turned off; fundamental tones at 10MHz and 10.05MHz, third order distortion at frequencies of 9.95MHz and 10.1MHz. Results measured using test setup of Fig. 5.10.

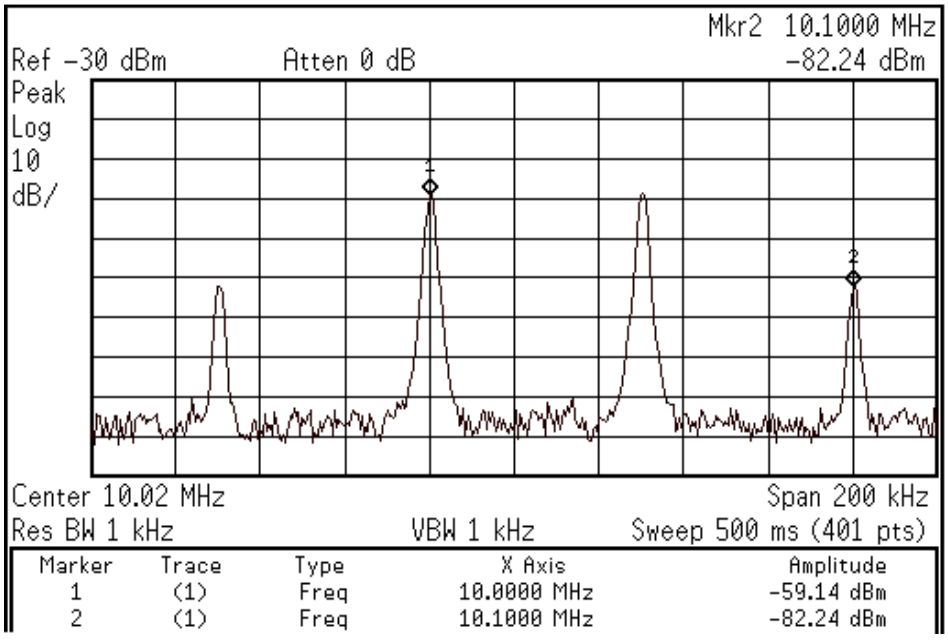

Figure 5.20 Measured output frequency spectrum of amplifier A2 at 10MHz: amplifier A1 is deactivated; fundamental tones at 10MHz and 10.05MHz, third order distortion at frequencies of 9.95MHz and 10.1MHz. Results measured using test setup of Fig. 5.10.

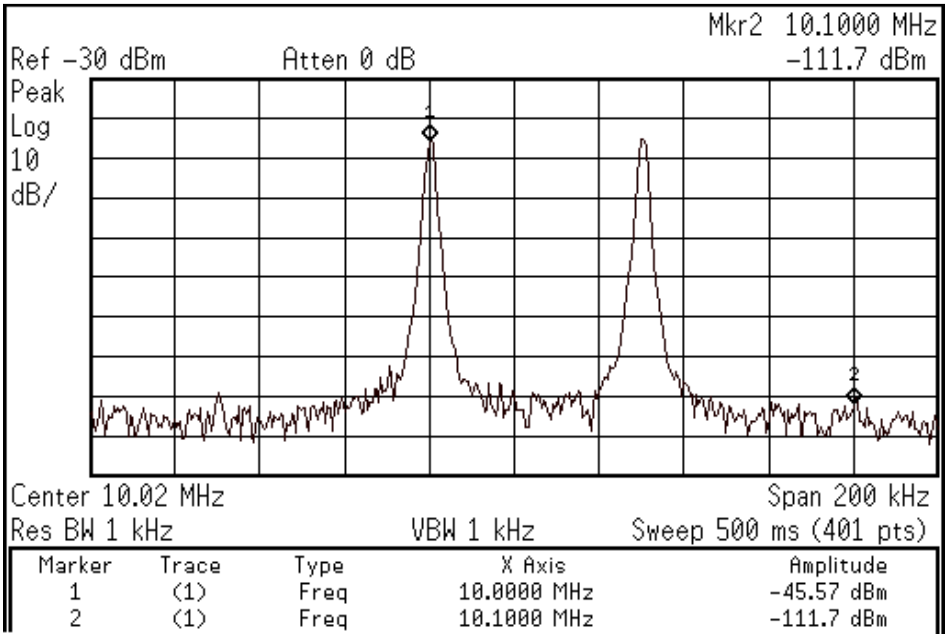

Figure 5.21 Measured output frequency spectrum of composite amplifier at 10MHz showing cancellation of third order distortion. Results measured using test setup of Fig. 5.10.

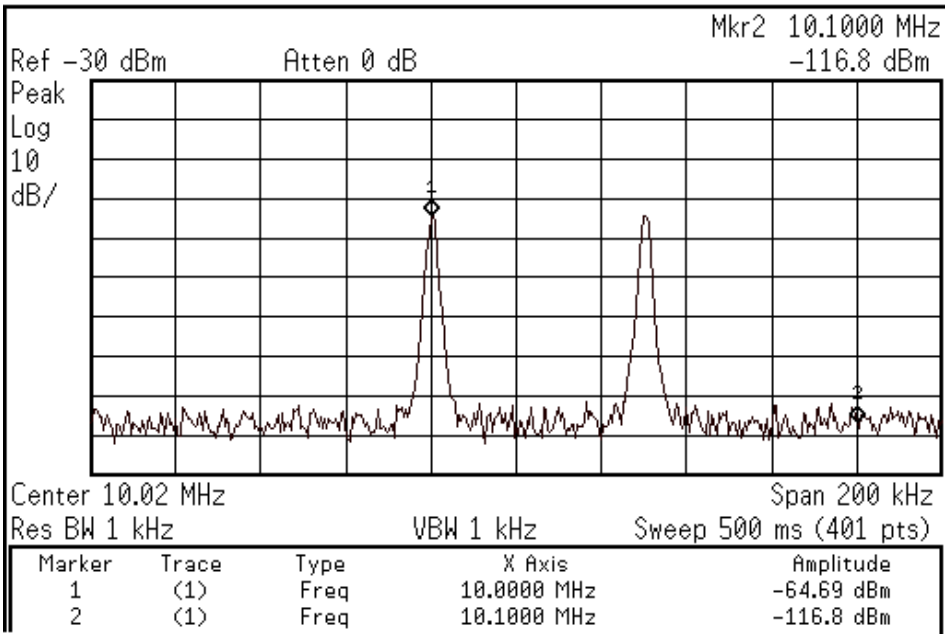

Figure 5.22 Measured signal leakage between input and output of test circuit at 10MHz; fundamental tones at frequencies 10MHz and 10.05MHz. Results measured using test setup of Fig. 5.10.

Table 5.6 Summary of measured results for 10 MHz using setup from Fig. 5.10

	P_{in} (dBm)	P_{out} (dBm)	G (dB)	P_{3rd} (dBm)	OIP3 (dBm)
Amplifier A1 (Fig. 5.19)	-10.1	-44.2	-34.1	-81.8	-25.4
Amplifier A2 (Fig. 5.20)	-10.1	-59.1	-49.0	-82.2	-47.6
Composite amplifier A1+A2 (Fig. 5.21)	-10.1	-45.6	-35.5	-111.7	-12.6
Comparison: A1 with comp. Amp			Gain loss: 1.4 dB	Third order suppression: 29.9 dB	Improvement of OIP3: 12.8 dB

Table 5.6 lists measured results for a test frequency of 10 MHz similar to Table 5.2. A suppression of third order distortion of approximately 30 dB is achieved resulting in an amelioration of 12.8 dB in OIP3. The composite amplifier suffers a gain loss of 1.4 dB. Again, the following calculations illustrate the validity of equation (4.7) using measured results from Table 5.6:

$2(OIP3_1-OIP3_2)=3(G_1-G_2)$

$2(OIP3_1-OIP3_2)=2(-25.4+47.6)=44.4dB$

$3(G_1-G_2)=3(-34.1+49.0)=44.7dB$

In the following presentation, a fourth and last set of frequency spectra for a test frequency of 30 MHz is discussed. Two sinusoidal signals of frequencies 30 MHz and 30.05 MHz are applied to the test circuit. The four frequency spectra of the set are shown in Figs. 5.23 to 5.26. The input power level at the input of the test circuit was kept constant at –10.1 dBm. First, the output spectrum of amplifier A1 is shown in Fig. 5.23.

The fundamental tones have a power level of –44.1 dBm. Third order intermodulation products at frequencies 29.95 MHz and 30.1 MHz show a power level of –81.1 dBm. Secondly, Fig. 5.24 depicts the output power level of amplifier A2. Third order distortion of A2, -81.7 dBm, has approximately the same power level as third order distortion of A1. Furthermore, the fundamental power levels have a value of about –53.6 dBm. The output signal of the composite amplifier exhibits the frequency spectrum as illustrated in Fig. 5.25. Third order intermodulation products are still visible but reduced to –105.6 dBm. The third order distortion suppression can be calculated as 24.5 dB. The fundamental tones show a power level of –45.2 dBm. Compared with the spectrum of A1, a reduction of gain of approximately 1 dB is noted. Finally, as previously stated, the leakage for a frequency of 30 MHz is shown in Fig. 5.26. The fundamental tones are shown with a leakage power level of –55.0 dBm. The results for a test frequency of 30 MHz are summarized in Table 5.7.

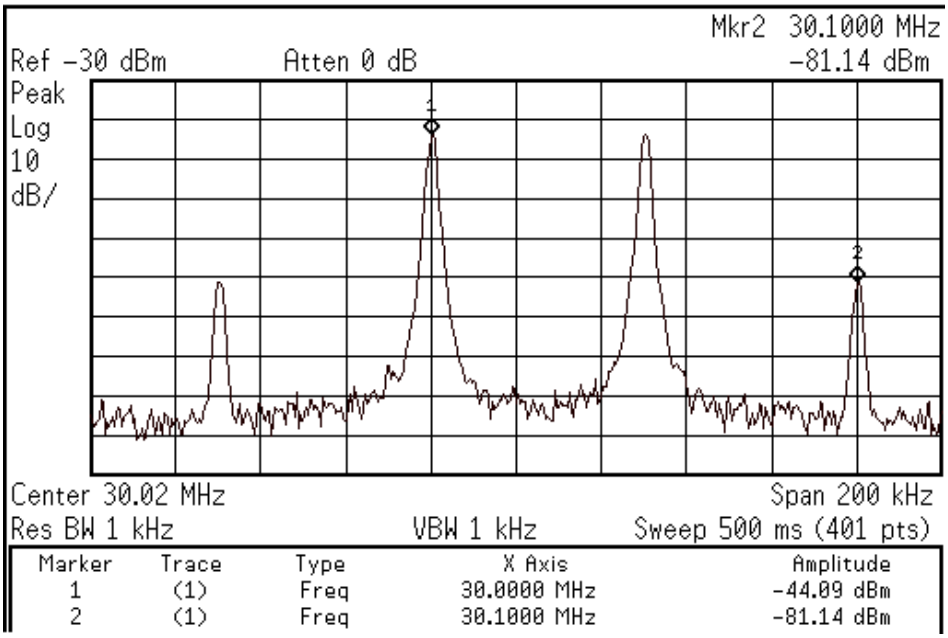

Figure 5.23 Measured output frequency spectrum of amplifier A1 at 30MHz: amplifier A2 is deactivated; fundamental tones at 30MHz and 30.05MHz, third order distortion at frequencies of 29.95MHz and 30.1MHz. Results measured using test setup of Fig. 5.10.

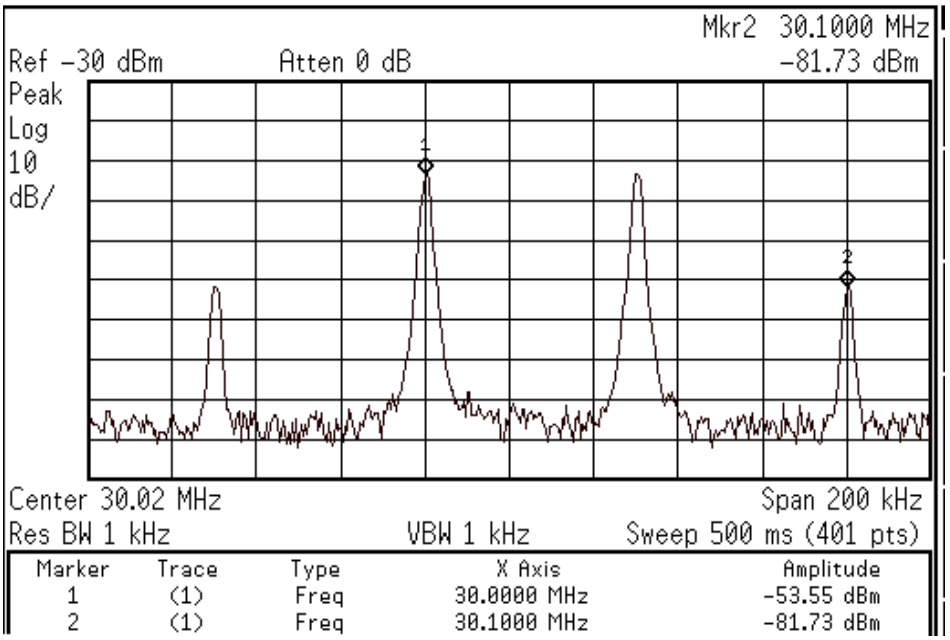

Figure 5.24 Measured output frequency spectrum of amplifier A2 at 30MHz: amplifier A2 is deactivated; fundamental tones at 30MHz and 30.05MHz, third order distortion at frequencies of 29.95MHz and 30.1MHz. Results measured using test setup of Fig. 5.10.

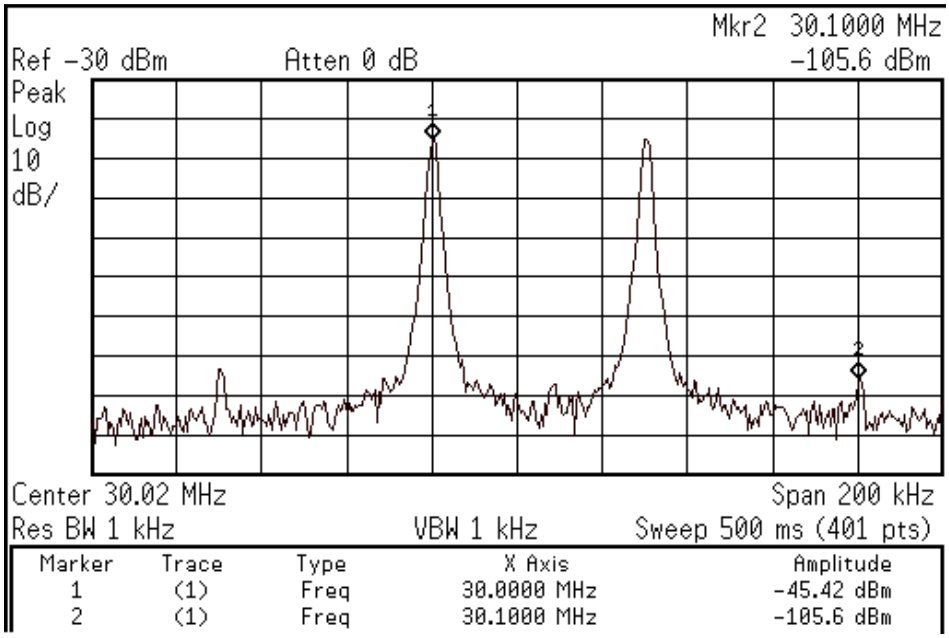

Figure 5.25 Measured output frequency spectrum of composite amplifier showing cancellation of third order distortion; third order distortion reduced to –105.6dBm at marker 2. Results measured using test setup of Fig. 5.10.

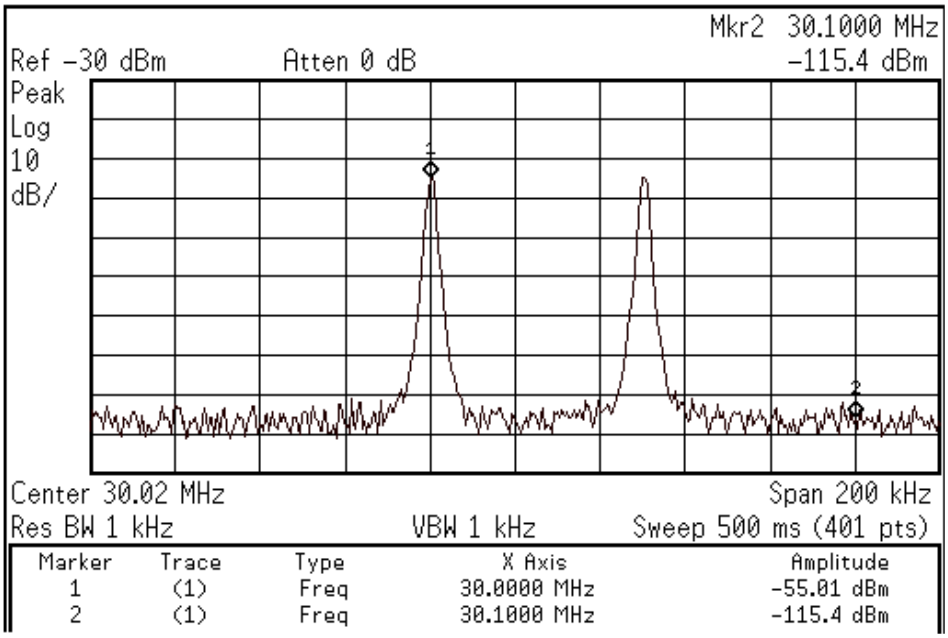

Figure 5.26 Measured signal leakage between input and output of test circuit at 30MHz; fundamental tones at frequencies 30MHz and 30.05MHz. Results measured using test setup of Fig. 5.10.

Table 5.7 Summary of measured results for 30 MHz using setup from Fig. 5.10

	P_{in} (dBm)	P_{out} (dBm)	G (dB)	P_{3rd} (dBm)	OIP3 (dBm)
Amplifier A1 (Fig. 5.23)	-10.1	-44.1	-34.0	-81.1	-25.6
Amplifier A2 (Fig. 5.24)	-10.1	-53.6	-43.5	-81.7	-39.6
Composite amplifier A1+A2 (Fig. 5.25)	-10.1	-45.4	-35.3	-105.6	-15.3
Comparison: A1 with comp amp.			Gain loss: 1.3 dB	Third order suppression: 24.5 dB	Improvement of OIP3: 10.3 dB

The measured results are summarized in Table 5.7 for a frequency of 30 MHz, similar to Table 5.2, indicate a suppression of third order distortion of 24.5 dB. As a result OIP3 is improved by 10.3 dB. A gain loss of 1.3 dB was measured. Equation (4.7) is proved true for a frequency of 30 MHz, shown in the following calculations using measured results and equation (4.7).

$2(OIP3_1-OIP3_2)=3(G_1-G_2)$

$2(OIP3_1-OIP3_2)=2(-25.6+39.6)=28.0dB$

$3(G_1-G_2)=3(-34.0+43.5)=28.5dB$

The results of the evaluation of equation (4.7) are almost equal, thus confirm the validity of (4.7) also for a test frequency of 30 MHz.

The variations in third order distortion suppression in dependence of frequency are listed in Table 5.8. Starting at a low frequency, suppression increases and reaches its highest value of approximately 30 dB for 10 MHz. For increasing frequencies,

Table 5.8 Summary of third order distortion suppression and improvement in OIP3 with variations in frequency

Frequency (MHz)	Third order suppression (dB)	Improvement in OIP3 (dB)	Reduction in Gain (dB)
1	25.1	12.9	1.4
3	28.9	12.3	1.4
10	29.9	12.8	1.4
30	24.5	10.3	1.3
110	22.6	8.5	2.0

suppression decreases and arrives at the lowest value of 22.6 dB at 110 MHz. One explanation for variations in suppression could be variations in gain of the composite amplifier within the measured bandwidth. Variations in gain can be created if the two amplifiers, A1 and A2, have ripples in their gain characteristic. Higher signal leakage at higher frequencies contributes most likely to the lower suppression at higher frequencies. Signal leakage could mask the output signal of the less linear amplifier A2.

5.4 Power Efficiency Considerations

For handheld devices, power consumption plays a major role since their battery lifetime depends on power consumption. The following section briefly discusses a few considerations concerning power efficiency. As calculated in Table 5.2, the third order output intercept point was improved by approximately 8.5 dB ($OIP3_c$-$OIP3_1$) with an input power level of –17 dBm at the input of the test circuit. Transferring the improvement level into a linear power ratio leads to the ratio of the possible new RF power to old RF power ($P_{RFnew}/P_{RFold} = 10^{0.85} = 7.08$) with a value of 7.08. In essence, the composite amplifier performs like an amplifier with 7 times the output power capability.

On the other hand, the auxiliary amplifier contributes to the DC power consumption with approximately 26 μA. The total DC current in the circuit adds up to 303 μA. Power consumption is defined as $P_{DC}=V_{DD}*I_{SS}$. Consequently, since the supply voltage remains constant, the ratio of new DC power to old DC power can be introduced. It is $P_{DCnew}/P_{DCold}=303\mu A/287\mu A=1.056$. This leads to one definition of power efficiency which is RF power (typically at the drain of a transistor) divided by DC power, P_{RF}/P_{DC} [1]. Now the ratio of the factors of RF power and DC power can be used to find the efficiency improvement, $7.08/1.056\approx6.7$, where factor 6.7 is the power efficiency improvement of the composite amplifier. In essence, a 1-Watt amplifier would have the third order performance of a 7.08 Watt amplifier at the cost of 1.056 increase in DC power. One way to view this situation is, the composite amplifier can deliver more RF power with higher power consumption (I_{SS2} = 26 μA), or take an amplifier with unaltered RF power with reduced DC power consumption; for instance, reduced bias current. The second way of viewing the result is very favorable for battery-powered handheld devices such as cellular phones because less power consumption extends battery life.

5.5 Summary of Measured Results

The presented measured results show that the simple algebraic expression (4.7) is valid over a wide frequency range from 1 MHz to more than 100 MHz. A reduction of third order intermodulation products of up to 29 dB can be achieved resulting in an improvement of 12.8 dB in third order output intercept point. Some third order suppression occurs above the 1dB compression point of the less linear amplifier. Cancellation of third order distortion holds up to the 1dB compression point of the less linear amplifier. However, cancellation of third order distortion degrades at the 1dB

compression point of the more linear amplifier because the two amplifiers compress. A maximum gain loss between more linear amplifier and composite amplifier of 2 dB was measured at a frequency of 110 MHz. Gain loss occurs due to subtraction of fundamental signals at the final output in Fig. 5.1.

Equation (5.4) as condition of elimination of third order distortion proposed in [21] does not hold for the prototype, possibly because of velocity saturation. Conversely, equation (4.7) is proved to be valid and does not depend on operating conditions of transistors.

The measured results indicate that the presented topology, as only one possible implementation of the new linearization method, is useful for applications with low input power levels such as radio receivers. To eliminate third order distortion up to the 1dB compression point of a power amplifier, a new topology is being investigated. This newer topology is presented in Chapter 7 and promises third order distortion cancellation up to the 1dB compression point of the higher-power amplifier.

CHAPTER 6: SIMULATION OF CROSS-COUPLED DIFFERENTIAL PAIR

Within Chapter 6, simulation results for the cross-coupled differential pair are discussed for a frequency of 110 MHz and constant input power. Furthermore, simulated, measured and calculated results are compared for various input power levels.

To gain information about the accuracy of simulation results regarding nonlinear behavior of integrated circuits, the cross-coupled differential pair in Fig. 5.1 was simulated using Agilent's Advanced Design System (ADS) [33] (The ADS schematic is in Appendix F). It would be advantageous to predict the nonlinear behavior of integrated circuits during the design phase using simulation tools. In addition, the accuracy of simulation results depends on accurate device models. Comparison of measured and simulated results can give information about the accuracy of transistor models.

ADS is capable of performing a harmonic balance simulation including nonlinear effects such as intermodulation and gain compression. The simulation was set-up similar to the test setup in Fig. 5.2. Two attenuators at the input are not necessary, since there is only one signal source. An amplifier and an additional attenuator are added to the output of the test circuit, to obtain the proper output power levels. The bias tee was not used in the simulation. The simulation was also performed single-ended. To achieve the most accurate results, transistor models for this particular process with lot number T2AK-AZ, were obtained from the fabrication service MOSIS [35]. The transistor models used are of type BSIM3 version 3. The BSIM3v3 transistor model card can be found in Appendix D.

Harmonic balance simulation allows the user to input two or more sinusoidal frequency tones into a circuit and calculate intermodulation products and harmonics. A two-tone test can be simulated using harmonic balance simulation. The test circuit was simulated with a two-tone test at frequencies of 110 MHz and 110.1 MHz.

First, the bias conditions of the cross-coupled differential pairs were varied in order to find the optimal condition for cancellation of third order distortion. The bias voltages at the gates of transistors M1 through M4 in Fig. 5.1 were kept constant at 4.5 V, whereas the bias currents were changed by modifying the gate voltages (V_{bias1}, V_{bias2} in Fig. 5.1) of transistors M5 and M6. The input power level of 0 dBm at the signal source was kept constant. Table 6.1 lists bias conditions for the measured prototype and simulation results under the condition of cancellation. The increased bias current of amplifier A2 in the simulation might be a result of inaccurate modeling of transistor M6 in Fig. 5.1. Transistor M6 is biased at the low end of strong inversion where the transistor does no longer exhibit square-law characteristic. The lower bias current could also be an effect of velocity saturation.

After the prerequisite for cancellation was established, the harmonic balance

Table 6.1 Comparison of bias conditions of prototype and simulation

	Amplifier A1		Amplifier A2	
	Prototype	Simulation	Prototype	Simulation
V_{bias}	1.06 V	0.91 V	0.867 V	0.874 V
I_{SS}	287 μA	287 μA	26 μA	86 μA

simulation provided the results shown in Figs. 6.1 through Fig. 6.3. The output power was taken at the final output of the simulation setup in order to be able to compare measured and simulated results. Fig. 6.1 depicts the output frequency spectrum of amplifier A1 when amplifier A2 was turned off. The fundamental spectral lines appear at 110 MHz and 110.1 MHz with a power level of –29.7 dBm and are shown as innermost spectral lines. Third order intermodulation products are created at frequencies 109.9 MHz and 110.2 MHz with power levels of –75.0 dBm, depicted in Fig. 6.1 as outermost spectral lines. The third order output intercept point can be calculated to –7.05 dBm. Next, the output frequency spectrum of amplifier A2 is illustrated in Fig. 6.2, with fundamental tones at 110 MHz and 110.1 MHz and power level of –38.0 dBm. Third order distortion products have a power level of –74.3 dBm at frequencies 109.9 MHz and 110.2 MHz. The third order intercept point of the less linear amplifier, A2, equals about –19.7 dBm. A third spectrum in Fig. 6.3, shows the frequency content of the output signal of the composite amplifier. A reduction of third order distortion is clearly visible. Third order distortion was reduced to –115 dBm. The fundamental tones have a power level of –34 dBm. Comparing the power levels of the fundamental tones in Fig. 6.1 and Fig. 6.3, a decrease in gain of approximately 4 dB can be noted.

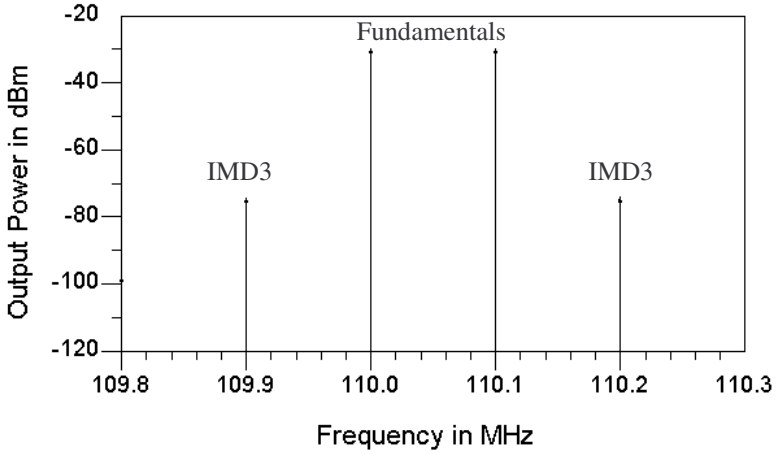

Figure 6.1 Simulated (ADS) output frequency spectrum of amplifier A1 at 110MHz: Amplifier A2 turned off; fundamental components are at 110MHz and 110.1MHz with power level of –29.7dBm, third order distortion at frequencies 109.9MHz and 110.2MHz and –75dBm power level. Results using circuit in Fig 5.1.

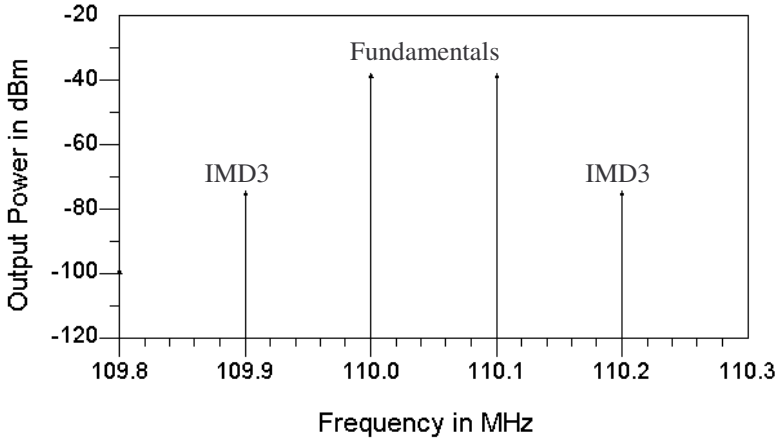

Figure 6.2 Simulated (ADS) output frequency spectrum of amplifier A2 at 110MHz: Amplifier A1 disabled; fundamental components are at 110MHz and 110.1MHz with power level of –38dBm, third order distortion at frequencies 109.9MHz and 110.2MHz and –74.3dBm power level. Results using circuit in Fig. 5.1.

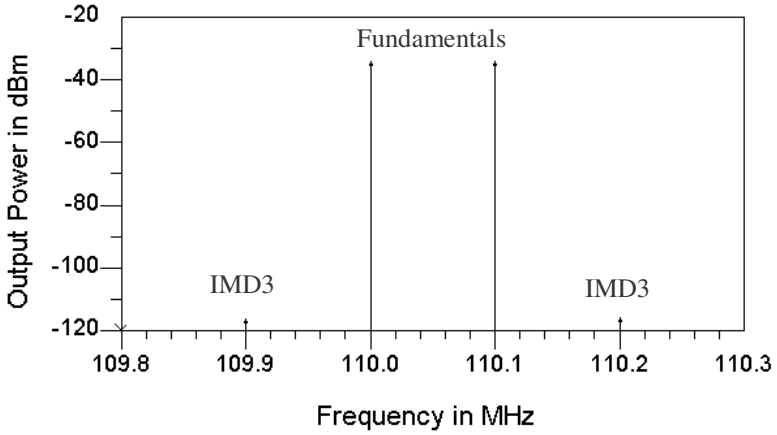

Figure 6.3 Simulated (ADS) output frequency spectrum of composite amplifier at 110MHz showing cancellation of third order distortion: third order products are reduced to –115 dBm. Results using circuit in Fig. 5.1.

The simulation results of Fig. 6.1, 6.2 and 6.3 are summarized in Table 6.2. Table 6.2 lists simulation results in the same manner as Table 5.2. Input power P_{in}, fundamental output power P_{out} and third order distortion levels P_{3rd} are listed and OIP3 and gain are calculated for amplifier A1, A2 and composite amplifier. An improvement in third order suppression of 40 dB was simulated, resulting in an increase in OIP3 of 14 dB. Due to higher bias current of amplifier A2, composite amplifier suffers a gain loss of 4 dB. Measured results for 110 MHz indicate only a loss of 2 dB in fundamental output power. The most likely reason is the increased transconductance of the less linear amplifier g_{mA2} in the case of simulation due to the higher bias current of 86 μA. The voltage gain of the cross-coupled differential pair is described with $(g_{mA1}-g_{mA2})R_D$ (5.1). An increase in transconductance g_{mA2} results in lower gain of the composite amplifier. Applying calculated OIP3 and gain to equation (4.7) proves its validity.

Table 6.2 Summary of simulation results for 110 MHz using schematic in Appendix F

	P_{in} (dBm) (at signal source)	P_{out} (dBm) (at final output)	Gain (dB)	P_{3rd} (dBm)	OIP3 (dBm)
Amplifier A1 (Fig. 6.1)	0	-29.7	-29.7	-75.0	-7.1
Amplifier A2 (Fig. 6.2)	0	-38.0	-38.0	-74.3	-19.7
Composite amplifier A1+A2 (Fig. 6.3)	0	-34.0	-34.0	-115.0	6.9
Comparison: A1 with compostie amp.			Gain loss: 4.3 dB	Third order suppression: 40 dB	Improvement of OIP3: 14 dB

The following calculations show prove of equation (4.7) using simulation results from Table 6.2. Substituting simulated results into (4.7) leads to

$2(OIP3_1-OIP3_2)=3(G_1-G_2)$

$2(OIP3_1-OIP3_2)=2(-7.05+19.73)=25.4dBm$

$3(G_1-G_2)=3(-29.7+38.0)=24.9dB$

As one can see, equation (4.7) also proves to be a useful equation describing the cancellation condition very well.

In addition to the simulations with constant input power, harmonic balance simulations were performed with varying input power. The main interest lies in the behavior of third order cancellation between the 1dB compression points of less and more linear amplifiers. One can expect that for low input power levels the condition for cancellation described with equation (4.7) holds and that third order suppression

gradually begins to decrease for input power levels greater than the 1dB compression point of the less linear amplifier.

The input power level was varied from –10 dBm to +10 dBm. Fig. 6.4 through Fig. 6.6 display the simulated linear fundamental output power of amplifier A1, amplifier A2 and composite amplifier together with measured data from Chapter 5 and calculated results from Chapter 4. The input power was taken from the signal source and not the input of the test circuit. Similar, the final output power of the ADS schematic shown in Appendix F was recorded. The plot in Fig. 6.4 depicts linear output power over input power of amplifier A1 with amplifier A2 deactivated. The solid curve depicts calculated data, the curve marked with triangles shows the measured data in Fig. 6.4, and simulated data is marked with squares. One can see that for low input power levels up to approximately 4 dBm input power, calculated, measured and simulated data, agree very well. For input power levels above 4 dBm, simulated data deviates from measured and calculated data. Calculated data does not have a gradual clipping behavior due to the definition of output power in Mathcad. The simulated data has a 1dB input compression point of about 6 dBm, approximately 3 dB lower than the P1dB (9 dBm) of measured data.

Next, a comparison of simulated, measured and calculated output power of amplifier A2 is presented in Fig.6.5 similar to Fig. 6.4. Again, the solid line presents the calculated data. Measured data is marked with triangles in Fig. 6.5. For lower input power, measured and calculated data agree very well. Conversely, simulated output power level, marked with squares in Fig. 6.5, is about 2 dB higher. The reason for the increased gain is higher bias current, hence higher gain for amplifier A2. A 1dB input

compression point of about 4 dBm was found. Finally, Fig. 6.6 depicts the linear output power of the composite amplifiers for simulated, measured and calculated data. The solid line characterizes the calculated output power. As before, measured and calculated data agree well. Simulated data has a 1dB input compression point of about 6 dBm. As expected in Fig. 6.6, simulated output power levels are approximately 3 dB lower than measured and calculated data because of the greater bias current in amplifier A2, thus greater transconductance g_{mA2}. All simulated data in Fig. 6.4 through Fig 6.6 show less linear behavior compared to measured data for high input power levels. A possible reason could be that the test circuit becomes strongly nonlinear for high input power levels. As a result, transistor models and simulation tool lose accuracy in calculating the power levels.

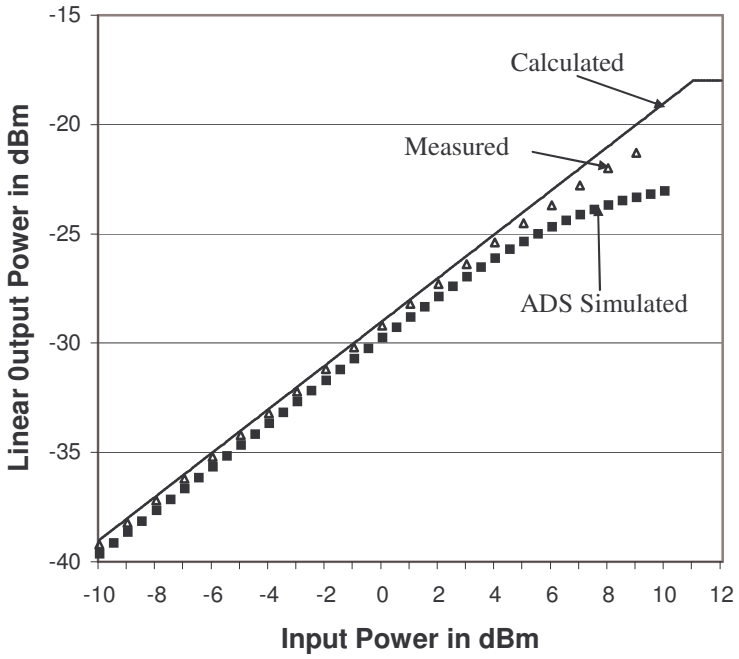

Figure 6.4 Simulated, measured and calculated linear output power of amplifier A1 at 110MHz: Amplifier A2 is turned off. Simulated results using ADS schematic.

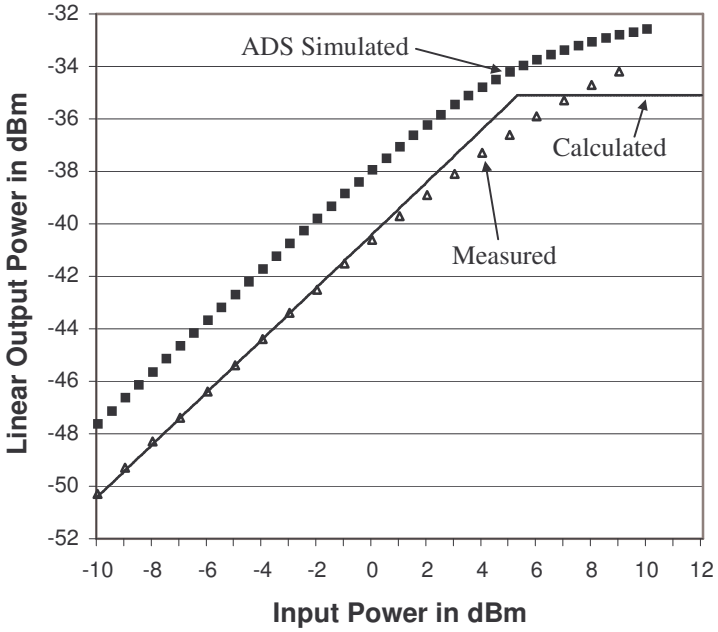

Figure 6.5 Simulated, measured and calculated linear output power of amplifier A2 at 110MHz: Amplifier A1 is turned off. Simulated results using ADS schematic.

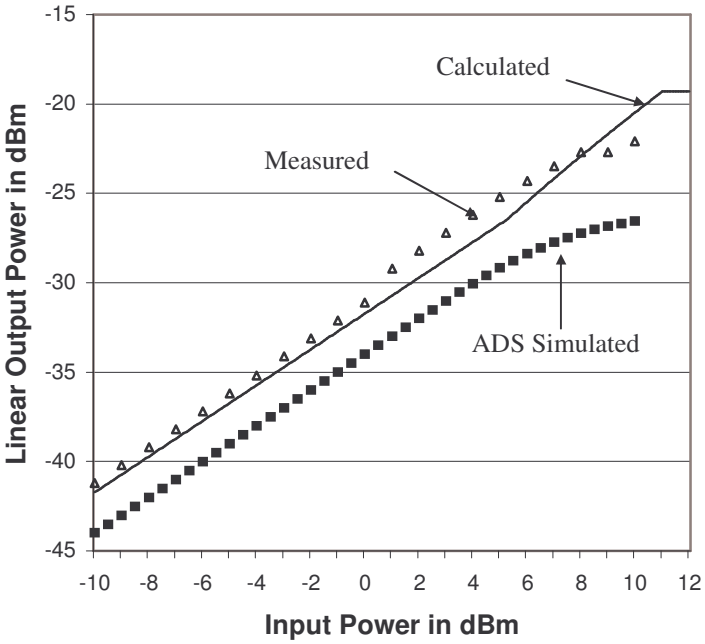

Figure 6.6 Simulated, measured and calculated linear output power of composite amplifier at 110MHz: Amplifiers A1 and A2 are turned on for cancellation of third order distortion. Simulated results using ADS schematic.

A comparison of third order intermodulation products of simulated, measured and calculated data is illustrated in Fig. 6.7 through Fig. 6.9. The plots are separated into results of amplifier A1, A2 and composite amplifier similar to Figs 6.4 to 6.6 and the traces are marked in the same manner. First, Fig. 6.7 displays third order distortion of amplifier A1 when amplifier A2 is turned off. Simulated data, illustrated with squares in Fig. 6.6, deviates from measured data and exhibits higher distortion levels at about 4 dBm at higher input power levels. Measured data is shown in Fig. 6.6 as triangles. The solid line in Fig. 6.7 shows the calculated data from Chapter 4. Secondly, third order distortion of amplifier A2 only is depicted in Fig. 6.8, where amplifier A1 is deactivated. Again, simulation of amplifier A2 reveals higher third order distortion for input power levels greater than 3 dBm compared to measured data. Calculated data from Chapter 4 is included in Fig. 6.8 as a solid line. Finally, Fig. 6.9 depicts third order distortion versus input power levels with simulated, measured and calculated data. Fig. 6.9 is also showing cancellation of third order intermodulation products for low input power levels. The characteristic of calculated data in Fig. 6.9 is shown as the solid curve. For higher input power levels, third order intermodulation products rise rapidly because third order distortion of amplifier A1 and amplifier A2 are not equal and no longer cancel. As before, simulated data has higher third order intermodulation products for input power levels greater than 4 dBm compared to measured data. The calculated data in Fig. 6.9 exhibits ideal cancellation of third order distortion for input power levels below the 1dB compression point of amplifier A2.

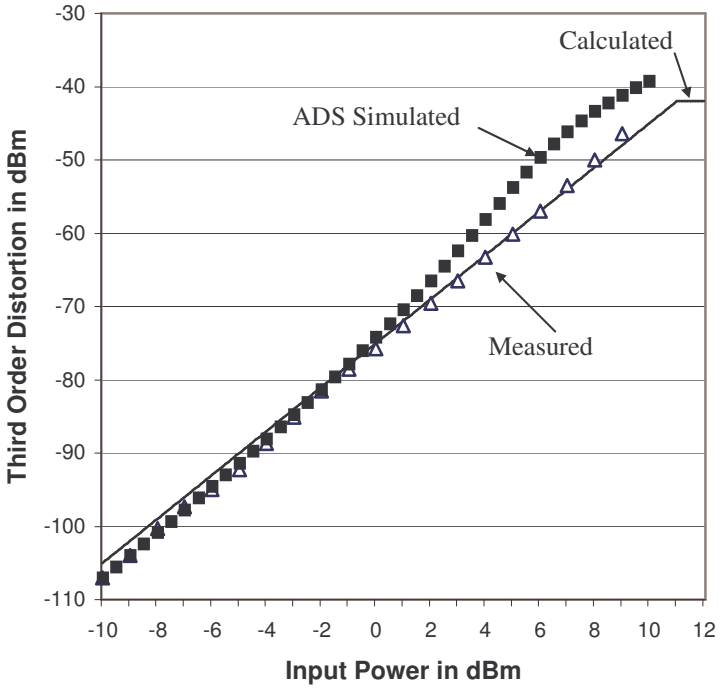

Figure 6.7 Simulated, measured and calculated third order distortion of amplifier A1 at 110MHz: Amplifier A2 is turned off. Simulated results using ADS schematic.

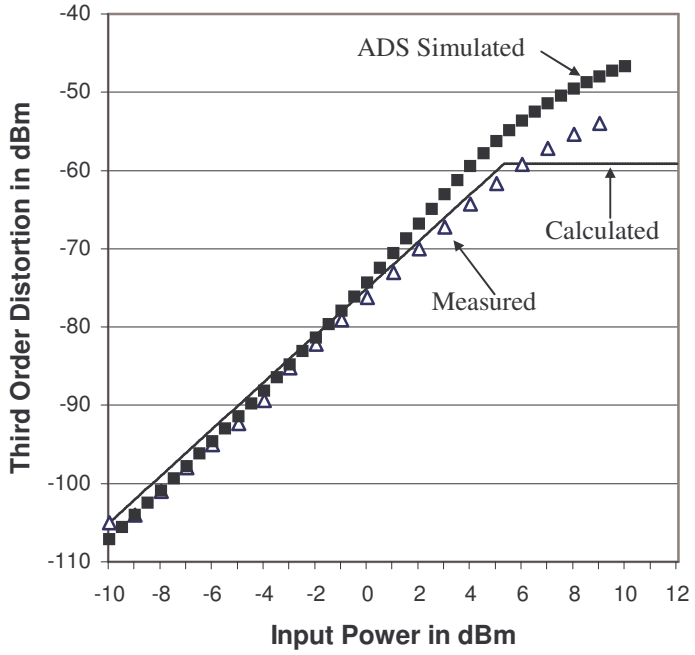

Figure 6.8 Simulated, measured and calculated third order distortion of amplifier A2 at 110MHz: Amplifier A1 deactivated. Simulated results using ADS schematic.

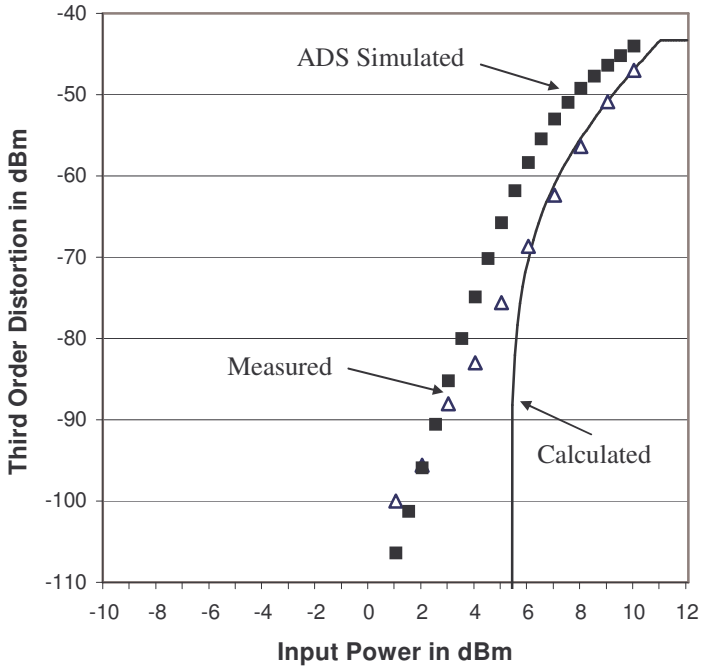

Figure 6.9 Simulated, measured and calculated third order distortion of composite amplifier at 110MHz: Amplifiers A1 and A2 are activated for cancellation of third order distortion. Simulated results using ADS schematic.

In conclusion, simulation results of fundamental output power in Fig. 6.4 through Fig. 6.6 exhibit more compression than the measured results do, which indicates that the simulated circuit is more nonlinear than the measured prototype. As a result of the increased nonlinearity of simulated data, third order distortion is also increased for the simulation results in Fig. 6.7 through Fig. 6.9. The deviation of simulated and measured results is most likely a result of the inability of the simulation tool to accurately model the compressed integrated circuit.

CHAPTER 7: CONCLUSION

This chapter summarizes and concludes the work presented in this thesis. First, a summary of the work from previous chapters is given. The second section suggests future work. The second section also presents a different topology with four amplifier stages applying the new linearization method. This topology promises good linearization up to the saturation point of the main amplifier.

7.1 Summary of Work

This work introduced new methods of linearization. The new linearization techniques are based on an unusual concept: a good amplifier combined with a bad amplifier results in a better amplifier. In this context, a good amplifier is a more linear amplifier and a bad amplifier is a less linear amplifier. A basic topology, Fig. 4.1, was presented using two amplifiers connected in parallel where the output signals are subtracted. With proper adjustment of amplifier parameters, such as gain and third order output intercept point OIP3, nonlinear components of the final output signal can be eliminated and linear components ideally are not affected. Theoretical conditions for cancellation of third order distortion in terms of power series coefficients and in terms of gain and third order output intercept point were derived. The derived equation (4.7) for cancellation of third order distortion, $2(OIP3_1-OIP3_2)=3(G_1-G_2)$ is especially useful in practice where gain and OIP3 of amplifiers is available. Due to this simple concept, the

new linearization techniques lend themselves to integrated circuit implementation. Some of the advantages of the new linearization methods are no need for external expensive parts, low complexity and small size. Mathematical considerations in Chapter 4 showed that this particular topology achieved good results in cancellation of third order intermodulation products for low input power levels up to the 1dB compression point of the less linear amplifier.

Chapter 2 showed an overview of prior linearization techniques such as feedback linearization, feedforward linearization and predistortion. Most of these linearization methods are fairly complex or suffer from device matching restrictions and feedback limitations. Many of the prior linearization techniques are only applicable to transmitter and not receiver designs. Conversely, the new linearization technique can also be used for receivers.

In Chapter 3, effects of nonlinearity were presented. Desensitization occurs when a strong interferer and a weak desired signal are processed simultaneously by a radio. The radio receiver becomes less sensitive to the desired signal. Cross modulation was discussed as a second effect of nonlinearity. In cross modulation, the modulation of a carrier is imposed on the adjacent carrier. Finally, third order intermodulation was shown to be problematic because third order intermodulation products appear close to the desired band of interest and are difficult to filter. Intermodulation occurs when fundamental and harmonics interact and create new, so called, intermodulation products. Furthermore, two common measures of nonlinearity were presented in Chapter 3. Output intercept point and 1dB compression point describe a device's nonlinear behavior.

The cross-coupled differential pair was presented in Chapter 5. This particular

circuit is fully integrable and was utilized as a first prototype. Previous research showed that third order distortion cancellation can be expressed in terms of bias currents and aspect ratios of transistors (5.4). The cross-coupled differential pair was implemented in a 0.5 μm CMOS process. Measurement and simulation results revealed that good elimination of distortion occurs at low input power levels (up to 30 dB improvement) when no compression occurs in the amplifiers, however, some improvement of linearity can still be seen between compression points of less linear and more linear amplifiers. Additional measurements for different frequencies showed that third order distortion cancellation holds for a frequency range of 1 MHz to 110 MHz for the prototype. Furthermore, an increase in power efficiency by a factor of 6 due to improvement in linearity could be achieved.

Finally, simulation results presented in Chapter 6 agree very well with the measured results for low input power levels. Simulated data shows less linear behavior of linear output power and higher third order distortion levels for high input power levels compared to measured results. This characteristic of simulated data suggests that the simulation tool, ADS, is not able to simulate the saturated amplifier accurately.

7.2 Future Research

In the following discussion, Future Work is presented. Future research is necessary to bring the new linearization method to product level. To apply the new linearization technique to power amplifiers, new topologies must be designed to account for linearization up to the 1dB compression point of the PA. Such a topology is illustrated in Fig. 7.1 using four amplifier stages. The circuit in Fig. 7.1 consists of four amplifiers, one attenuator and two subtractors. On the left side, the input signal is equally distributed

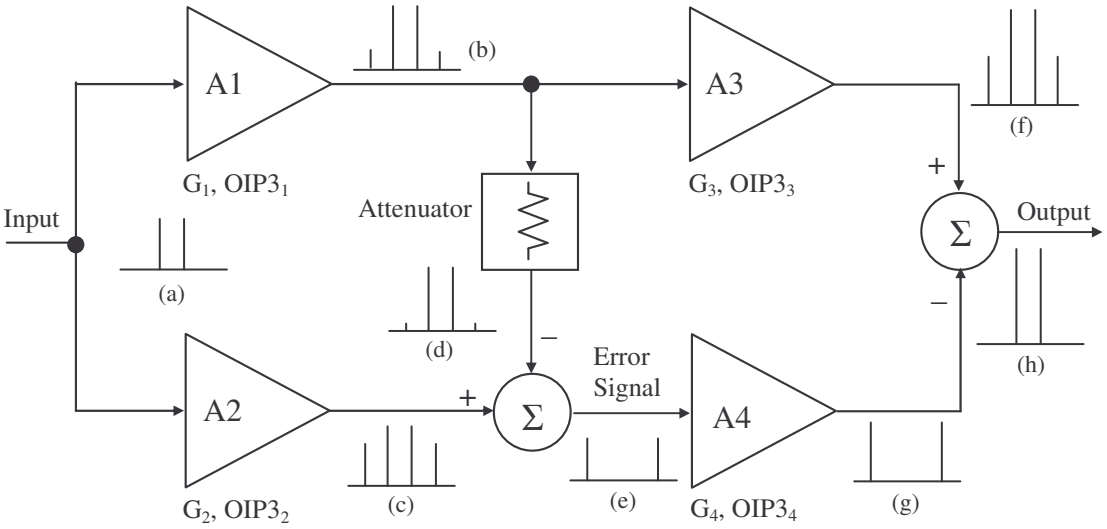

Figure 7.1 Four-stage amplifier linearization: A1, A2, A3 and A4 are amplifiers with parameters G and OIP3. After attenuation by an attenuator, the output signal of A1 is subtracted from output signal of A2. The error signal is amplified by amplifier A4 and subtracted from output signal of amplifier A3. A3 is the power amplifier. The attenuator is typically a passive network. The frequency spectrum of every path is shown in (a) to (h).

between amplifier A1 and amplifier A2. The output signal of amplifier A1 is attenuated and input to a subtractor. The second input to the subtractor is connected to the output of amplifier A2. The part described so far is called the first loop, which cancels (ideally) the carrier. In Fig. 7.1, the output of amplifier A1 is also connected to the input of amplifier A3. The output signal of the first subtractor is called error signal and is inputted to amplifier A4. The error signal ideally only contains third order distortion. Output signals of amplifiers A3 and A4 are subtracted with a second subtractor at the output of the circuit, thus eliminating third order distortion. The last described mechanism in the second loop ideally cancels third order distortion. Fig. 7.1(a) to Fig. 7.1(h) show frequency spectra of every signal path in the circuit. A two-tone test signal is applied to the input of amplifiers A1 and A2 shown in Fig. 7.1(a). Fig. 7.1(b) depicts the output spectrum of amplifier A1 consisting of fundamental tones and third order distortion.

Similarly, frequency spectrum (c) in Fig. 7.1 shows the frequency content of the output of amplifier A2 with higher third order distortion products than amplifier A1. The spectrum in Fig. 7.1(d) is an attenuated version of spectrum (b). After subtraction, the linear signals are eliminated (ideally), as shown in Fig.7.1(e). The error signal is amplified, resulting in frequency spectrum Fig. 7.1(g). The output of amplifier A3 has a spectrum, depicted in Fig. 7.1(f), with increased amplitude of fundamentals and third order distortion. If third order distortion levels in (f) and (g) have equal amplitude, the output signal has a frequency spectrum as illustrated in Fig. 7.1(h), showing elimination of third order distortion. Amplifier A1 needs to be fairly linear. Conversely, amplifier A2 should be designed with higher distortion levels. Amplifier A3 is the power amplifier to be linearized and should be the first amplifier to clip. Error amplifier A4 must be designed for high linearity to prevent creation of more distortion. Next, theoretical considerations are performed to gain insight into the performance of the topology in Fig. 7.1.

Each of the amplifiers in Fig. 7.1 can be described with gain and third order output intercept point OIP3. In order to gain information about the characteristic and the level of third order distortion suppression, Mathcad was used to perform preliminary calculations using gain, G, and OIP3.

For the following calculations, some assumptions have to be made. First, the linear output power for all amplifiers was defined as input power plus gain, P_{in}+G. When the amplifier reaches compression, the output power level is defined as OIP3-12dBm. The 1dB compression point is typically about 12 dB lower than the third order output intercept point, as was discussed in Chapter 4. Since each single amplifier has two independent parameters, it seems that there are 8 independent parameters with four

amplifiers. The opposite is the case. Some of the parameters are dependent on each other. To cancel the carrier in the first loop, two equal amplitudes have to be subtracted at the first subtractor in Fig. 7.1. In order to have equal amplitudes at the inputs of the subtractor, the following dependence exists: $G_2=G_1$-Att. The output signal of amplifier A1 has to be attenuated to reach the necessary power level. Amplifier A2 is required to be less linear, thus has lower gain than amplifier A1. As a rule of thumb, the gain difference between amplifier A1 and amplifier A2 is 10 dB. The parameters shown in Table 7.1 are independent and are chosen as listed. All parameters of all amplifiers can be calculated using independent parameters in Table 7.1. A summary of amplifier parameters is given in Table 7.2. The attenuation is 10 dB. The complete Mathcad program is attached in Appendix E.

Table 7.1 Independent parameters for four-stage amplifier linearization

Independent Parameters	Description (using Fig. 7.1)
G_1=20 dB	Gain of amplifier A1
$OIP3_1$=10 dBm	Third order output intercept point of amplifier A1
G_3=30 dB	Gain of amplifier A3
$Deltaoip_{12}$=10 dB	Difference in output intercept point of amplifiers A1 and A2, derated to inputs of first subtractor
Oip3overdrive=10 dB	oip3overdrive=($OIP3_1$+G_3)–$OIP3_3$

Table 7.2 Summary of gain and third order intercept point of four-stage amplifier linearization from Fig. 7.1

A1		A2		A3		A4	
G_1	$OIP3_1$	G_2	$OIP3_2$	G_3	$OIP3_3$	G_4	$OIP3_4$
20 dB	10 dBm	10 dB	-10 dBm	30 dB	30 dBm	41.7 dBm	10 dBm

Of interest is the behavior of the linearization of amplifier A3 that is operated around the 1dB compression point. To exercise the amplifier in the critical operating region, the input power was swept from –40 to 0 dBm. Fig. 7.2 depicts third order distortion suppression and linear output power of amplifier A3 as a function of input power. The suppression is defined as the difference between third order distortion of the final output minus third order distortion of amplifier A3. The linear output power of A3 reaches the compression point at an input power level of –32 dBm. The solid line in Fig. 7.2 illustrates the characteristic of third order suppression. Suppression is greater than 40 dB up to the compression point of the power amplifier A3. The topology in Fig. 7.1 promises elimination of third order distortion at the 1dB compression point of the power amplifier. The third order distortion characteristics at the final output are shown in Fig. 7.3. For low input power levels up to –32 dBm, third order intermodulation products of A3 and A4 are equal and cancel at the final output. Above the 1dB compression point of amplifier A3, all distortion levels reach a final maximum and clip.

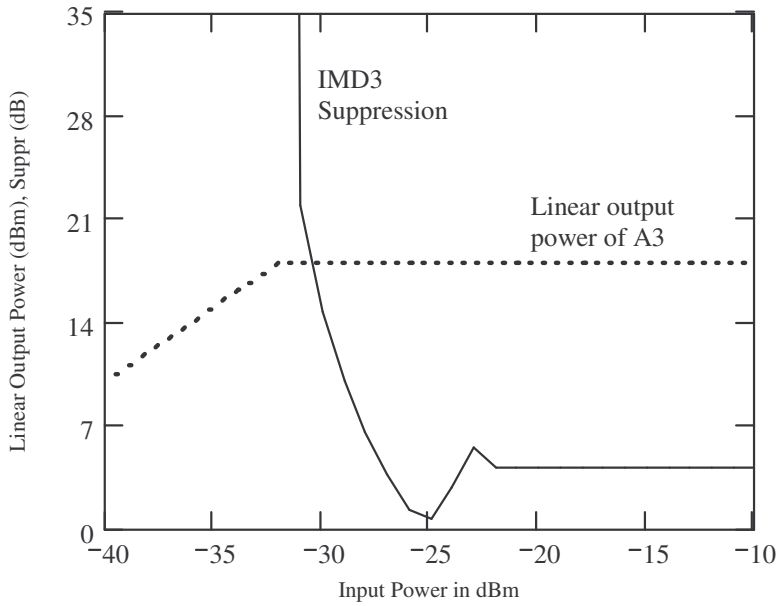

Figure 7.2 Simulated (Mathcad) third order distortion suppression of four-stage amplifier linearization: linear output power of amplifier A3 (dotted line) and third order distortion suppression at the final output (solid line). Suppression up to 1dB compression point of A3. Linear power compresses at –32dBm

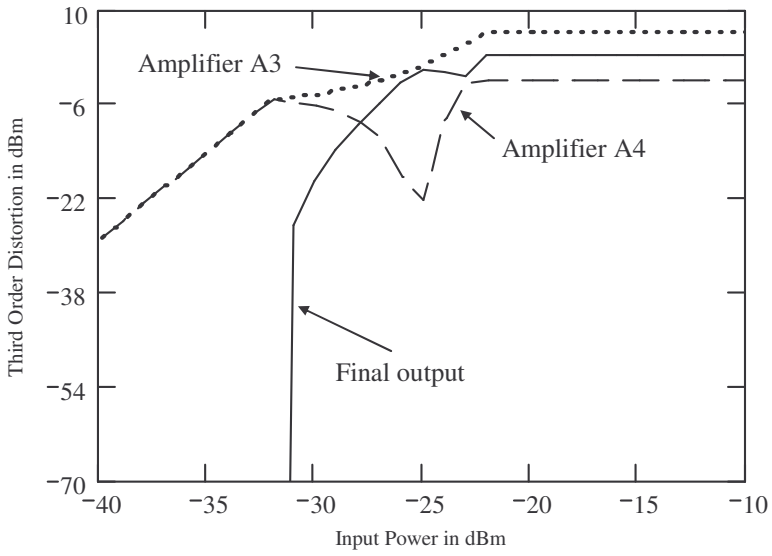

Figure 7.3 Simulated (Mathcad) third order distortion cancellation of four-stage amplifier linearization: dotted line shows IMD3 of amplifier A3 and dashed line illustrates IMD3 of error amplifier A4, solid line depicts cancellation of third order distortion at final output up to a input power level of –31dBm.

One possible circuit topology for the four-stage approach is shown in Fig. 7.4. The circuit in Fig. 7.4 is based on differential pairs and has a differential input voltage and differential output voltage. Amplifier A1 is comprised of MOS transistors M1 and M2 and amplifier A2 consists of transistors M3 and M4. Transistors M1 to M5 built a cross-coupled differential pair in Fig.7.4 to achieve subtraction and create an error signal. Additional resistors in the drains of transistors M1 and M2 correspond to the attenuator in Fig. 7.1. The main signal can be taken from the drains of transistors M1 and M2 whereas the error signal is available at the cross-coupled drains. Furthermore, amplifier A3 is comprised of MOS transistors M5 and M6 forming a differential pair and amplifier A4 is built using transistors M7 and M8. The subtraction at the output is also achieved by cross-coupling the drains of transistors M5 to M8. For a prototype, the current sources of the differential pairs need to be implemented using transistors.

So far only few preliminary harmonic balance simulations have been performed using ADS for the four-stage linearization. The bias currents of the transistors were adjusted to achieve reduction in fundamentals in the error signal and reduction of third order distortion at the final output. Preliminary calculations suggest that the transconductance g_m of amplifiers A1 and A2 have to be equal to cancel the linear fundamental components. Transconductance of amplifiers A3 and A4 must not be equal to prevent cancellation of the linear components. Table 7.3 lists the bias currents and aspect ratios of the transistors in Fig. 7.4 used for the simulation with ADS. These bias currents and aspect ratios achieve good results for the simulation, but further simulations have to be performed to find an optimum in bias conditions for a desired reduction of third order intermodulation distortion.

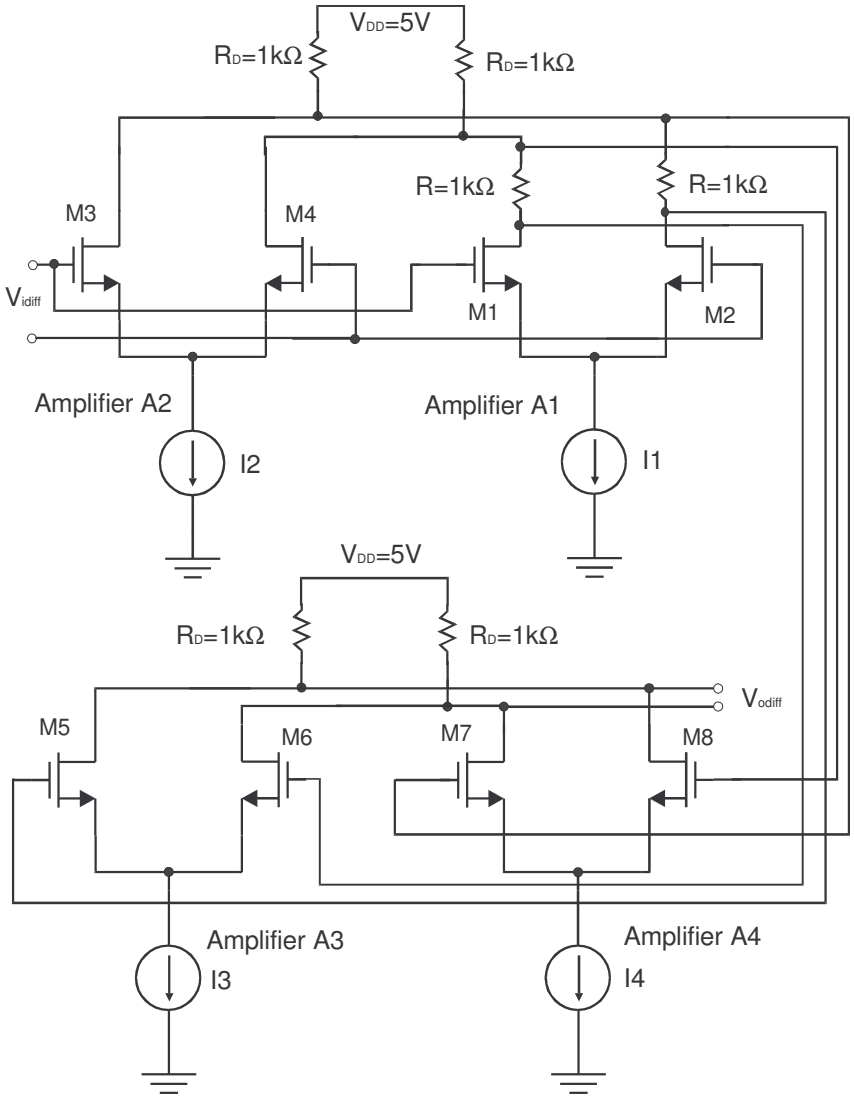

Figure 7.4 Topology for four-stage amplifier linearization using differential pairs. Amplifier A1 and A2 have cross-coupled drains and Amplifier A3 and A4 have cross-coupled drains. Amplifier A4 is the error amplifier. (The schematic is conceptual and omits decoupling circuits and biasing detail.)

The results of a harmonic balance simulation with varying input power levels form –40 to 0 dBm are shown in Fig. 7.5 to 7.7. First, Fig. 7.5 depicts the linear output power and the third order distortion of amplifier A1 in Fig. 7.4 over the input power. The upper curve displays the linear output power as solid trace. The trace below shows the third order distortion product as dash-dotted curve. As one can see, the linear output power of amplifier A1 does not compress, whereas IMD3 compresses. This phenomenon could not be explained yet.

Secondly, Fig. 7.6 illustrates the characteristic of the error signal over the input power as output of the subtractor. Again, the upper trace shows the linear component as solid curve and the dash-dotted curve displays the third order distortion. The linear component rises slowly with increasing input power. The reason for this behavior is the cancellation mechanism of the fundamental components. Furthermore, Fig. 7.6 indicates that the third order distortion product is not affected by the cancellation mechanism.

Lastly, Fig. 7.7 depicts the final output signal of Fig. 7.4 over input power. The upper trace displayed as solid curve, is the linear output power when amplifiers A2 and A4 are deactivated, then amplifiers A1 and A3 are cascaded and no subtraction occurs. This result is recorded as reference for the calculation of improvement of third order distortion. Equally, the third curve from the top is the third order distortion of cascaded amplifiers A1 and A3 when amplifier A2 and A4 are turned off. If cancellation is activated, then the four stages show a linear output power depicted as dotted curve at the top of Fig. 7.7. The fourth trace from the top shows third order intermodulation distortion when cancellation is turned on. In addition, the vertical left line marks the 1dB compression point of the cascaded amplifier A1 and A3 when A2 and A4 are turned off.

P1dB has a value of –20 dBm at an input power of –18 dBm. The 1dB compression point of the complete linearized amplifier also has a value of approximately –20 dBm but at an input power level of about –11 dBm. Furthermore, third order distortion is suppressed by approximately 30 dB up to the 1dB compression point of cascaded amplifiers A1 and A3 when A2 and A4 are deactivated. This result was suggested from the mathematical results of Mathcad. However, a gain loss of about 7 dB was simulated. This problem should be easily solvable by adjusting the proper parameters of the circuit in Fig. 7.4.

In summary, a third order distortion suppression of 30 dB was achieved resulting in an improvement in OIP3 of 7.75 dB. This improvement was reached at the 1dB compression point when amplifiers A2 and A4 were deactivated.

Table 7.3 Summary of bias currents and aspect ratios for four-stage amplifier linearization

	Bias current I	Aspect ratio W/L
Amplifier A1	1 mA	15
Amplifier A2	93 μA	60
Amplifier A3	160 μA	200
Amplifier A4	99 μA	1000

For future work it is suggested to continue optimizing the circuit in Fig. 7.4. In addition, other possible circuit topologies for the four-stage amplifier linearization should be considered differentially and single-ended.

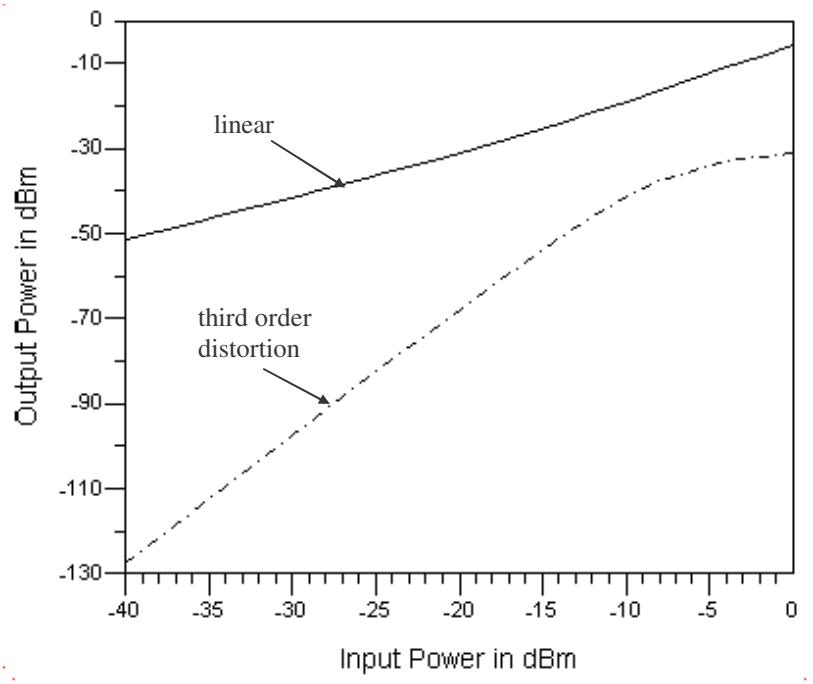

Figure 7.5 Simulated (ADS) linear and third order distortion output power of amplifier A1 using four-stage linearization from Fig.7.4 at 100 MHz.

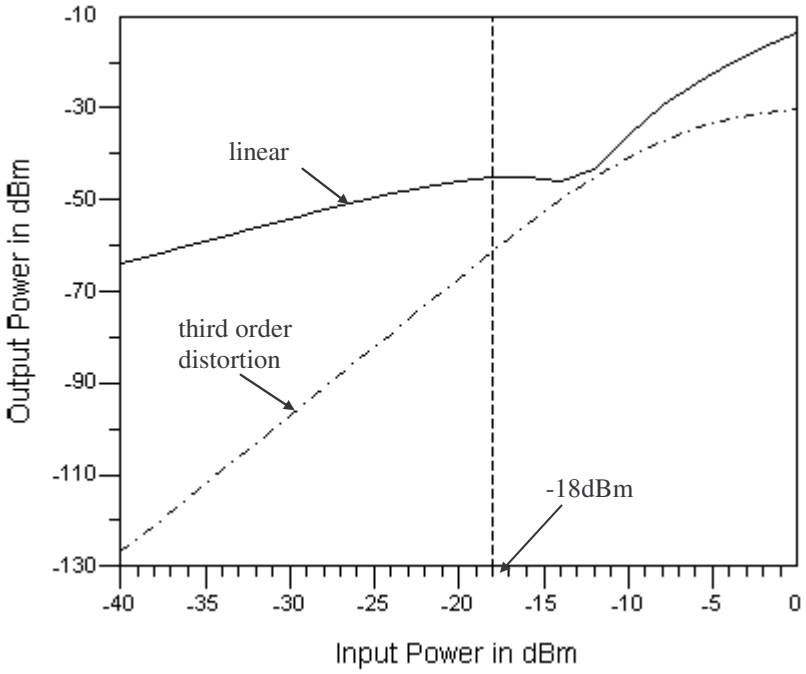

Figure 7.6 Simulated (ADS) linear power and third order distortion of error signal showing reduction of linear component: solid trace is linear component and dash-dotted curve is third order distortion. Results using Fig. 7.4 at 100 MHz.

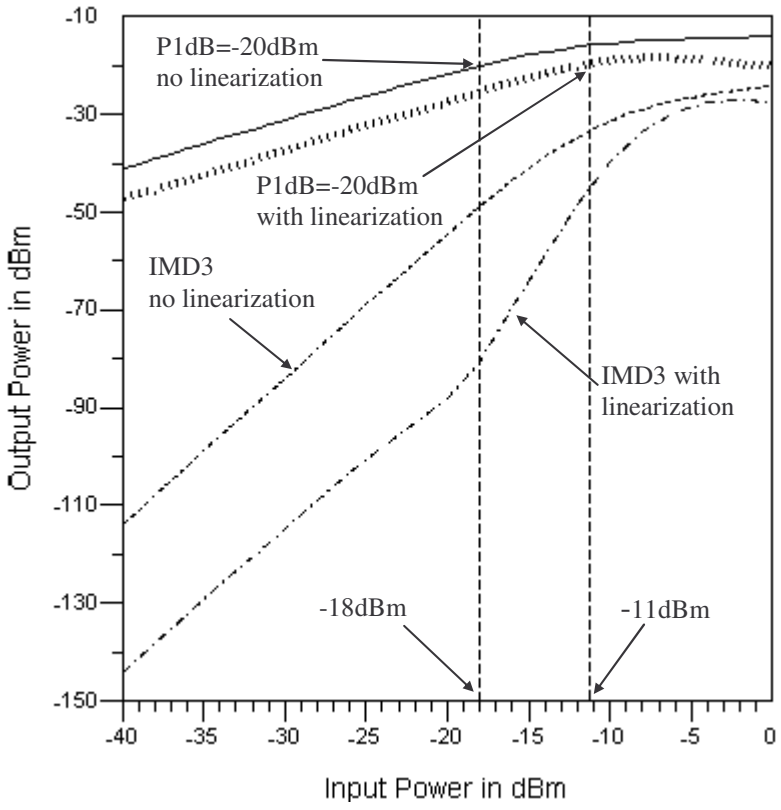

Figure 7.7 Simulated (ADS) linear output power and third order distortion of final output from Fig. 7.4 showing third order suppression: upper two curves depict linear output power, lower traces show third order intermodulation distortion; P1dB of circuit from Fig. 7.4 without linearization is –20dBm at –18 dBm input power, at 100 MHz.

So far, only differential circuits were implemented utilizing the new linearization method because of the simple subtraction mechanism. Most applications in radio systems, such as receiver and transmitter, require single-ended designs. Consequently, the subtraction mechanism has to be designed in a single-ended fashion.

Another recommendation is that research work should be put into investigating the effect of phase distortion. Many modern modulation schemes incorporate phase modulation, hence they need accurate phase reconstruction of the amplified signal. In the case at hand, both amplifiers exhibit AM/PM conversion perhaps because of different transistor sizes, thus different parasitic capacitances. Particularly pn-junctions are present in almost every device in integrated circuit design. pn-junctions have nonlinear voltage-dependent parasitic capacitances.

The downscaling of transistor sizes to submicron regions continues and modifies the square-law I-V characteristic of MOS devices. Short channel effects, such as velocity saturation and mobility reduction, arise and change the square-law relation to a more linear characteristic. This will affect the new linearization method, thus it is recommended that research should be done in this area.

The presented circuit topology was implemented in a 0.5 μm CMOS process. It is also of great interest to investigate the nonlinear behavior in different process technologies, such as 0.25 μm CMOS process, silicon-germanium BiCMOS process, gallium-arsenide or a process using silicon-on-insulator technology.

REFERENCES

[1] P. B. Kenington, *High-Linearity RF Amplifier Design*, Artech House, 2000

[2] B. Razavi, *RF Microelectronics*, Prentice Hall, 1998

[3] F. H. Raab et al, "Power Amplifiers and Transmitters for RF and Microwave," *IEEE Transactions on Microwave Theory and Techniques*, vol. 50, no. 3, March 2002

[4] A. Katz, "Linearization: Reducing Distortion in Power Amplifiers," *IEEE Microwave Magazine*, vol. 2, no. 4, December 2001

[5] P. B. Kenington, "Methods Linearize RF Transmitters And Power Amps (Part 1)," *Microwaves & RF*, vol. 37, no. 13, December 1998, pp. 102-116

[6] V. Petrovic, W. Gosling, "Polar Loop Transmitter," *IEE Electronics Letter*, vol. 15, no.10, May 1979, pp. 286-288

[7] M. Faulkner, "An automatic phase adjustment scheme for RF and Cartesian feedback linearizers," *IEEE Transaction on Vehicular Technology*, vol.49, May 2000, pp. 956-964

[8] J. L. Dawson, T. H. Lee, "Automatic phase alignment for high bandwidth Cartesian feedback power amplifiers," *IEEE Radio and Wireless Conference, 2000, RAWCON 2, 2000*, pp. 71-74

[9] P. B. Kenington, "Methods Linearize RF Transmitters and Power Amps (Part 2)," *Microwaves & RF*, vol. 38, no. 1, January 1999, pp. 79-89

[10] H. S. Black, *U. S. Patent 2,102,671*, Issued December 1937

[11] Y. Y. Woo, Y. Yang, J. Yi, J. Nam, J. Cha, B. Kim, "An Adaptive Feedforward Amplifier For WCDMA Base Stations Using Imperfect Signal Cancellation," *Microwave Journal*, April 2003

[12] C. B. Haskins, "Diode Predistortion Linearization for Power Amplifier RFICs in Digital Radios," *Thesis*, Virginia Polytechnic Institute and State University, April 2000

[13] S. P. Stapleton, "Amplifier Linearization Using Adaptive Digital Predistortion," *Applied Microwave & Wireless*, February 2001, pp. 72-77

[14] F. H. Raab, B. E. Sigmon, R. G. Myers, R. M. Jackson, "L-Band Transmitter Using Kahn EER Technique," *IEEE Transactions on Microwave Theory and Technology*, vol. 46, no. 12, December 1998

[15] D. K. Su, W. J. McFarland, "An IC for Linearizing RF Power Amplifiers Using Envelope Elimination and Restoration," *IEEE Journal of Solid-State Cicuits*, vol. 33, no. 12, December 1998

[16] D. C. Cox, "Linear amplification with nonlinear components," *IEEE Transactions on Communications*, vol. COM-22, December 1974, pp. 1942-1945

[17] A. Bateman, "The combined analogue locked loop universal modulator (CALLUM)," *Proceedings of the 422nd IEEE Vehicular Technology Conference*, May 1992, pp. 759-763

[18] Y. Yang, B. Kim, "A New Linear Amplifier Using Low-Frequency Second-Order Intermodulation Component Feedforwarding," *IEEE Microwave and Guided Wave Letters*, vol. 9, no. 10, October 1999

[19] Y. Ding, R. Harjani, "A +18dBm IIP3 in 0.35μm CMOS," International Solid-State Circuits Conference 2001, *Digest of Technical Papers, 2001*, pp. 162-163

[20] D. R. Webster, G. Ataei, D. G. Haigh, "Low-Distortion MMIC Power Amplifier Using s New Form of Derivative Superposition," *IEEE Transactions on Microwave Theory and Techniques*, vol. 49, no. 2, February, 2001

[21] C. Toumazou, F. J. Lidgey, D. G. Haigh, *Analogue IC design: the current-mode approach*, Peter Peregrinus Ltd., 1990

[22] J. Pedro, J. Perez, "An MMIC Linearized Amplifier using Active Feedback," *IEEE MTT-S*, Atlanta, Georgia, vol. 1, June 1993, pp. 95-98

[23] P. B. Kenington, "Linearized RF Transmitter Techniques," *IEEE MTT-S International Microwave Symposium, Tutorial Session on Advances in Amplifier Linearization*, Baltimore, MD, June, 1998

[24] S. W. Kim, H. Y. Cho, Y. Kim, I. S. Chang, "Design of a Predistorter Controlling Individual Orders of Intermodulation Using a New Harmonic Generator," *Microwave Journal*, April 2003

[25] B. Shi, L. Sundstroem, "A 200-MHz IF BiCMOS Signal Component Separator for Linear LINC Transmitters," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 7, July 2000

[26] S. A. Maas, *Nonlinear Microwave Circuits*, Artech House, 1988

[27] L. W. Couch II, *Digital and Analog Communication Systems*, Prentice Hall, New Jersey, 2001

[28] P. Wambacq, W. Sansen, *Distortion Analysis of Analog Integrated Circuits*, Kluwer Academic Publishers, 1998

[29] T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, Cambridge University Press, 1998

[30] T. P. Weldon, "Method and Apparatus for Cancellation of Third Order Intermodulation Distortion and Other Nonlinearities," *US Patent Application 10/101,005*, March 2002

[31] T. P. Weldon, K. Miehle, "Using Amplifiers with Poor Linearity to Linearize Amplifiers with Good Linearity," *IEEE International Microwave Symposium*, Philadelphia, PA, June 2003

[32] Mathsoft Engineering & Education, Inc, 101 Main Street, Cambridge, MA 02142, USA, http://www.mathcad.com

[33] Agilent Headquarters, 395 Page Hill Road, P.O. Box 10395, Palo Alto, California, 94303, USA, http://eesof.tm.agilent.com/

[34] P. R. Gray, P. J. Hurst, S. H. Lewis, R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, John Wiley & Sons, Inc, 2001

[35] The MOSIS Service, 4676 Admiralty Way, Marina del Rey, California, 90292-6695, USA, http://www.mosis.org

[36] MiniCircuits, P.O. Box 350166, Brooklyn, NY 11235, USA, http://www.minicircuits.com

[37] H. Khorramabadi, P. R. Gray, "High-Frequency CMOS Continuous-Time Filters," *IEEE Journal of Solid-State Circuits*, vol. Sc-19, no. 6, December 1984

APPENDIX A: POWER SERIES EXPANSION

For all following power series expansions common trigonometric identities are used. The input signal is denoted x(t), the output signal is denoted y(t). A is the amplitude of the input signal. c1, c2 and c3 are linear, second and third order power series coefficients, respectively. The frequency in radians per second is denoted ω_0.

A.1 Second Order Nonlinearity

$$x(t) = A\sin(\omega_0 t)$$
$$y(t) = c_1 x(t) + c_2 x^2(t) = c_1 A\sin(\omega_0 t) + c_2 A^2 \sin^2(\omega_0 t)$$
$$= \frac{1}{2} c_2 A^2 + c_1 A\sin(\omega_0 t) - \frac{1}{2} c_2 A^2 \cos(2\omega_0 t)$$

A.2 Third Order Nonlinearity

$$x(t) = A\sin(\omega_0 t)$$
$$y(t) = c_1 x(t) + c_3 x(t) = c_1 A\sin(\omega_0 t) + c_3 A^3 \sin^3(\omega_0 t)$$
$$= (c_1 A + \frac{3}{4} c_3 A^3)\sin(\omega_0 t) - \frac{1}{4} c_3 A^3 \sin(3\omega_0 t)$$

A.3 Second and Third Order Nonlinearities

$$x(t) = A\sin(\omega_0 t)$$
$$y(t) = c_1 x(t) + c_2 x^2(t) + c_3 x^3(t) = c_1 A\sin(\omega_0 t) + c_2 A^2 \sin^2(\omega_0 t) + c_3 A^3 \sin^3(\omega_0 t)$$
$$= \frac{1}{2} c_2 A^2 + (c_1 A + \frac{3}{4} c_3 A^3)\sin(\omega_0 t) - \frac{1}{2} c_2 A^2 \cos(2\omega_0 t) - \frac{1}{4} c_3 A^3 \sin(3\omega_0 t)$$

A.4 Second and Third Order Nonlinearities with Two-Tone Test Signal

$$x(t) = A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t)$$

$$y(t) = c_1 x(t) + c_2 x^2(t) + c_3 x^3(t)$$

$$= c_1 (A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t)) + c_2 (A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t))^2 + c_3 (A_1 \cos(\omega_1 t) + A_2 \cos(\omega_2 t))^3$$

$$= \frac{1}{2} c_2 A_1^{\,2} + \frac{1}{2} c_2 A_2^{\,2} \qquad \text{DC component}$$

$$+ (c_1 A_1 + \frac{3}{4} c_3 A_1^{\,3} + \frac{3}{2} c_3 A_1 A_2^{\,2}) \cos(\omega_1 t) + (c_1 A_2 + \frac{3}{4} c_3 A_2^{\,3} + \frac{3}{2} c_3 A_1^{\,2} A_2) \cos(\omega_2 t) \qquad \text{linear}$$

$$+ \frac{1}{2} c_2 A_1^{\,2} \cos(2\omega_1 t) + \frac{1}{2} c_2 A_2^{\,2} \cos(2\omega_2 t) \qquad \text{2nd harmonics}$$

$$+ \frac{1}{4} c_3 A_1^{\,3} \cos(3\omega_1 t) + \frac{1}{4} c_3 A_2^{\,3} \cos(3\omega_2 t) \qquad \text{3rd harmonics}$$

$$+ c_2 A_1 A_2 \cos((\omega_1 + \omega_2) t) + c_2 A_1 A_2 \cos((\omega_2 - \omega_1) t) \qquad \text{2nd order intermodulation}$$

$$+ \frac{3}{4} c_3 A_1^{\,2} A_2 \cos((2\omega_1 + \omega_2) t) + \frac{3}{4} c_3 A_1^{\,2} A_2 \cos((2\omega_1 - \omega_2) t) \qquad \text{3rd order intermodulation}$$

$$+ \frac{3}{4} c_3 A_1 A_2^{\,2} \cos((\omega_1 + 2\omega_2) t) + \frac{3}{4} c_3 A_1 A_2^{\,2} \cos((2\omega_2 - \omega_1) t) \qquad \text{3rd order intermodulation}$$

For $A_1 = A_2 = A$, y(t) simplifies to:

$$y(t) = c_2 A^2 \qquad \text{DC component}$$

$$+ (c_1 A + \frac{9}{4} c_3 A^3) \cos(\omega_1 t) + (c_1 A + \frac{9}{4} c_3 A^3) \cos(\omega_2 t) \qquad \text{linear}$$

$$+ \frac{1}{2} c_2 A^2 \cos(2\omega_1 t) + \frac{1}{2} c_2 A^2 \cos(2\omega_2 t) \qquad \text{2nd Harmonic}$$

$$+ \frac{1}{4} c_3 A^3 \cos(3\omega_1 t) + \frac{1}{4} c_3 A^3 \cos(3\omega_2 t) \qquad \text{3rd Harmonic}$$

$$+ c_2 A^2 \cos((\omega_1 + \omega_2) t) + c_2 A^2 \cos((\omega_2 - \omega_1) t) \qquad \text{2nd order intermod product}$$

$$+ \frac{3}{4} c_3 A^3 \cos((2\omega_1 + \omega_2) t) + \frac{3}{4} c_3 A^3 \cos((2\omega_1 - \omega_2) t) \qquad \text{3rd order intermod product}$$

$$+ \frac{3}{4} c_3 A^3 \cos((\omega_1 + 2\omega_2) t) + \frac{3}{4} c_3 A^3 \cos((2\omega_2 - \omega_1) t) \qquad \text{3rd order intermod product}$$

APPENDIX B: MATHCAD PROGRAM FOR TWO-STAGE AMPLIFIER

ip3_2ampkm.mcp

Mathcad program to calculate third order suppression and gain compression using two amplifier stages.

Forrnulae

p1=linear siganl
p3=3rd order level
G=gain dB, OIP=3rd order output intercept dBm

General compressed gain output power level formula:

$$p1(Pin, G, OIP) := \min(Pin + G, OIP - 12)$$ p1=Output power including compression

General third order product output power level formula:

$$p3(Pin, G, OIP) := OIP - 3(OIP - p1(Pin, G, OIP))$$ p3=Third order IMD output

Gain difference

$$dg(g1, g2) := g1 - g2 \qquad g2(g1, dg) := g1 - dg$$

Calculation of $OIP3_2$

$$oip2(oip1, dg) := oip1 - \frac{3}{2}dg$$

Calculation of final linear output component p1

$$finalpl(pin, g1, oip1, g2, oip2) := 20\log\left(10^{\frac{p1(pin, g1, oip1)}{20}} - 10^{\frac{p1(pin, g2, oip2)}{20}}\right)$$

Calculation of final third order component p3

$$finalp3out(pin, g1, oip1, g2, oip2) := 10\cdot\log\left[\left(10^{\frac{p3(pin, g1, oip1)}{20}} - 10^{\frac{p3(pin, g2, oip2)}{20}}\right)^2 + 10^{-40}\right]$$

Due to cancellation, argument in log function is 0 for equal third order components, in order to prevent undefined add 10^{-40} means -400dB

Calculation of suppression at final output

$$\text{suppression}(pin, g1, oip1, g2, oip2) := p3(pin, g1, oip1) - \text{finalp3out}(pin, g1, oip1, g2, oip2)$$

Final output third order component is subracted from third order output of main amplifier A1

Calculation of final output third order interceptpoint

$$\text{finaloip3}(pin, g1, oip1, g2, oip2) := \frac{(\text{finalpl}(pin, g1, oip1, g2, oip2) - \text{finalp3out}(pin, g1, oip1, g2, oip2))}{2} + \text{finalpl}(pin, g1, oip1, g2, oip2) \ ...$$

Example calculation

Independent Parameter: G1, $OIP3_1$ and gain difference

$Pin := 0$ Input power

$G1 := -29$ Gain of amplifier A1

$dg := 11.4$ Gain difference

$OIP1 := -6$ OIP3 of amplifier A1

Calculation of dependent parameter: G2 and $OIP3_2$

$G2 := g2(G1, dg)$ Gain of amplifier A2

$OIP2 := oip2(OIP1, dg)$ OIP3 of amplifier A2

Summary of amplifier parameter:

$G1 = -29$ $G2 = -40.4$

$OIP1 = -6$ $OIP2 = -23.1$

Calculation of power levels at output of amplifiers and final output

$p1(Pin, G1, OIP1) = -29$

$p3(Pin, G1, OIP1) = -75$

$p1(Pin, G2, OIP2) = -40.4$

$p3(Pin, G2, OIP2) = -75$

$\text{finalpl}(Pin, G1, OIP1, G2, OIP2) = -31.723$

$\text{finalp3out}(Pin, G1, OIP1, G2, OIP2) = -400$

$\text{suppression}(Pin, G1, OIP1, G2, OIP2) = 325$

APPENDIX C: SMALL SIGNAL GAIN OF CROSS-COUPLED DIFFERENTIAL PAIR

The small signal gain of the cross-coupled differential pair is derived in the following part. The circuit is depicted in Fig. 5.1. The half-circuit method is used for the derivation [34].

The circuit in Fig. 5.1 can be separated into two half circuits, shown in Fig. C.1. The first half-circuit on the left side in Fig. C.1, consists of transistors M1 and M4 with drain resistance R_D. The input voltages are $v_{idiff}/2$ and $-v_{idiff}/2$. The second half-circuit on the right side of Fig. C.1 is comprised of transistors M2 and M3 with drain resistor R_D. Input voltages are denoted $-v_{idiff}/2$ and $v_{idiff}/2$. The output voltages are v_{o1} and v_{o2}. The differential input voltage equals $v_{idiff}=v_{idiff}/2 -(-v_{idiff}/2)$. The body-effect is not in effect because body and source of the transistors have the same voltage potential. For the derivation of the small signal gain, the half-circuit on the left can be transferred to the small signal circuit depicted in Fig. C.2.

The MOS transistor model is well known and can be found in the literature [28], [34]. Transconductance of transistor M1 is denoted g_{m1} and transconductance of transistor M4 is denoted g_{m4}. g_{ds} is the conductance of the transistors in case channel length modulation is present and $R_D=1/G_D$ is the drain resistance. The following part shows the derivation of output voltage v_{o2} as a function of v_{idiff}.

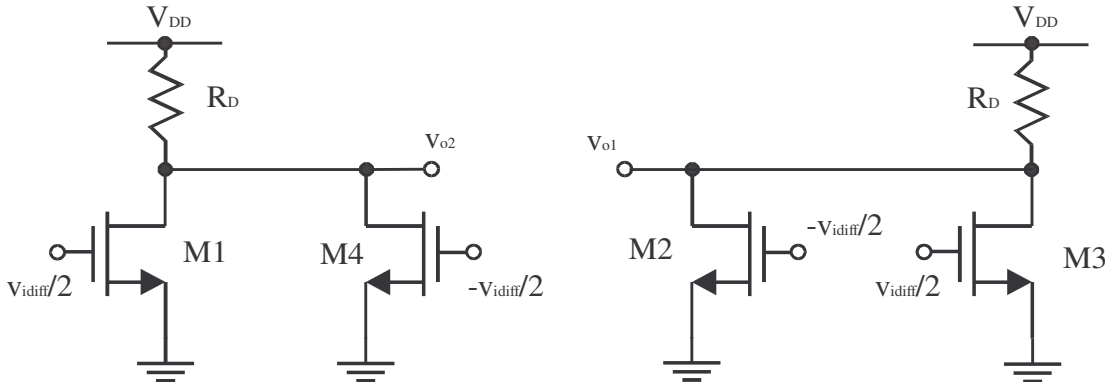

Figure C.1 Two half-circuits of cross-coupled differential pair: source terminals of all transistors are connected to virtual ground.

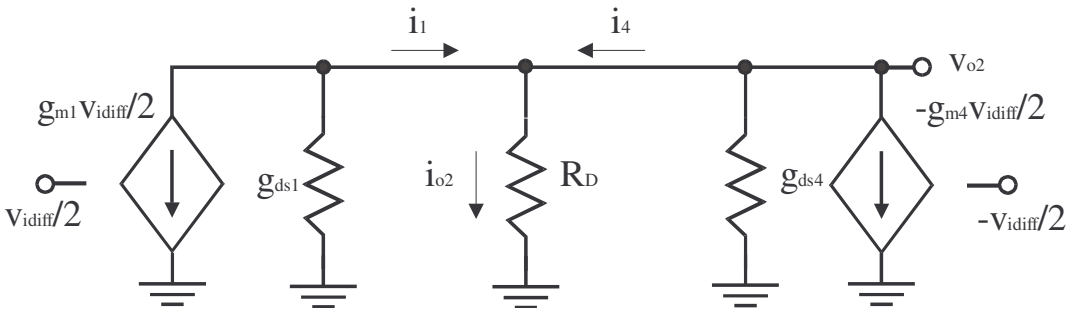

Figure C.2 Linearized small signal circuit for first half-circuit: on the left and right side are voltage controlled current sources, the currents of the current sources are $g_{m1}v_{idiff}/2$ and $-g_{m4}v_{idiff}/2$.

With the use of Kirchhoff's current law, the following equations can be derived:

$$i_1 = -v_{o2} g_{ds1} - g_{m1} \frac{v_{idiff}}{2} \tag{7.1}$$

$$i_4 = -v_{o2} g_{ds4} + g_{m4} \frac{v_{idiff}}{2} \tag{7.2}$$

$$i_{o2} = i_1 + i_4 = \frac{v_{o2}}{R_D} = -v_{o2} g_{ds1} - v_{o2} g_{ds4} - g_{m1} \frac{v_{idiff}}{2} + g_{m4} \frac{v_{idiff}}{2} \tag{7.3}$$

$$v_{o2} = -\frac{1}{2}(g_{m1} - g_{m4}) \frac{1}{g_{ds1} + g_{ds4} + G_D} v_{idiff} \tag{7.4}$$

The same derivations can be performed with the second half-circuit. Fig. C.3 illustrates the small signal circuit for the second half-circuit. Again, with the use of

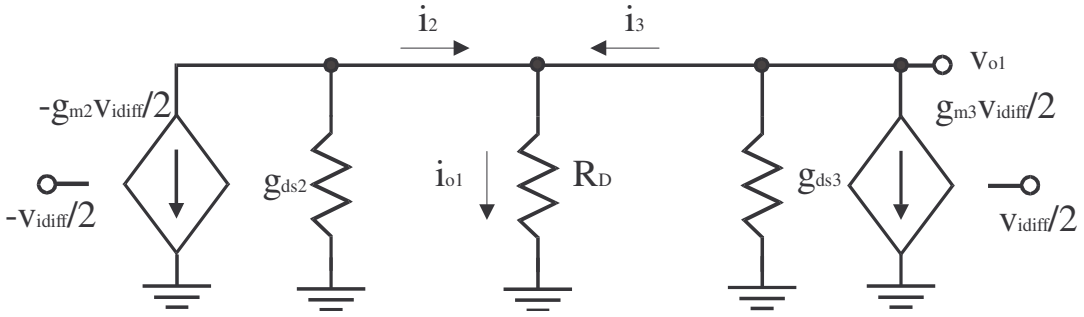

Figure C.3 Linearized small signal circuit for second half-circuit: on the left and right side are voltage controlled current sources, the currents of the current sources are $g_{m3}v_{idiff}/2$ and $-g_{m2}v_{idiff}/2$.

Kirchhoff's current law, the following equations can be found:

$$i_2 = -v_{o1}g_{ds2} + g_{m2}\frac{v_{idiff}}{2} \tag{7.5}$$

$$i_3 = -v_{o1}g_{ds3} - g_{m3}\frac{v_{idiff}}{2} \tag{7.6}$$

$$i_{o1} = i_2 + i_3 = \frac{v_{o1}}{R_D} = -v_{o1}g_{ds2} - v_{o1}g_{ds3} + g_{m2}\frac{v_{idiff}}{2} - g_{m3}\frac{v_{idiff}}{2} \tag{7.7}$$

$$v_{o1} = \frac{1}{2}(g_{m2} - g_{m3})\frac{1}{g_{ds2} + g_{ds3} + G_D}v_{idiff} \tag{7.8}$$

The differential output voltage is $v_{odiff}=v_{o1}-v_{o2}$. The following derivation shows the differential output voltage as a function of differential input voltage.

$$v_{odiff} = v_{o1} - v_{o2} = \left[\frac{1}{2}(g_{m2} - g_{m3})\frac{1}{g_{ds2} + g_{ds3} + G_D} + \frac{1}{2}(g_{m1} - g_{m4})\frac{1}{g_{ds1} + g_{ds4} + G_D}\right]v_{idiff} \tag{7.9}$$

for $g_{m1}=g_{m2}=g_{mA1}$ and $g_{m3}=g_{m4}=g_{mA2}$, the differential voltage gain can be found using (7.9).

$$A_{vdiff} = \frac{v_{odiff}}{v_{idiff}} = \frac{1}{2}(g_{mA1} - g_{mA2})\left(\frac{1}{g_1} + \frac{1}{g_2}\right) \tag{7.10}$$

where $g_1=g_{ds1}+g_{ds4}+G_D$ and $g_2=g_{ds2}+g_{ds3}+G_D$. If channel length modulation is neglected, then $g_{ds1}=g_{ds2}=g_{ds3}=g_{ds4}=0$ and (7.10) can be modified to equation (7.11)

$$A_{vdiff} = \frac{v_{odiff}}{v_{idiff}} = (g_{mA1} - g_{mA2})R_D \tag{7.11}$$

where A_{vdiff} is the differential voltage gain of the cross-coupled pair and v_{odiff} and v_{idiff} are the differential output and input voltages, respectively.

APPENDIX D: TEST PRINTED CIRCUIT BOARD LAYOUT

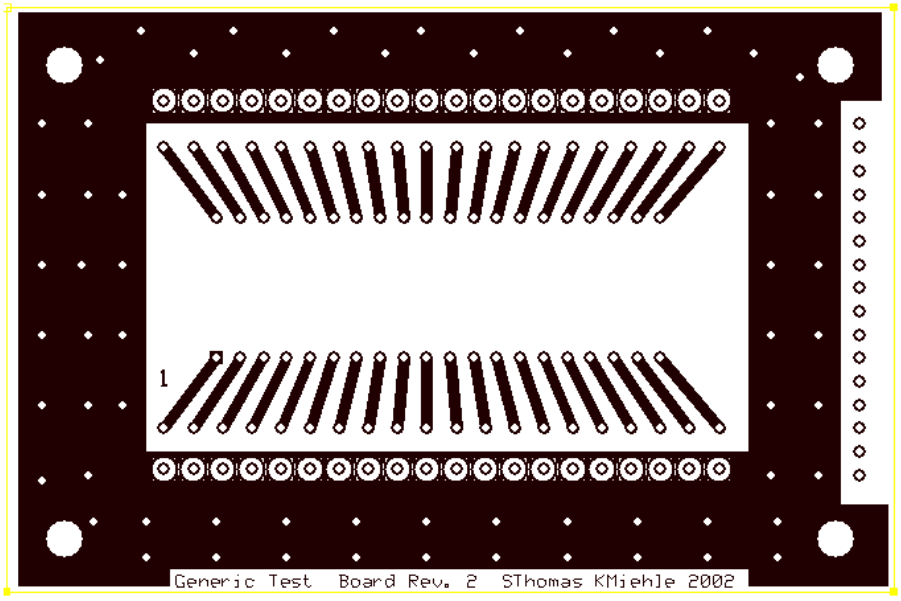

Figure D.1 Front side of printed circuit board used in test setup in Fig. 5.2 and Fig. 5.10.

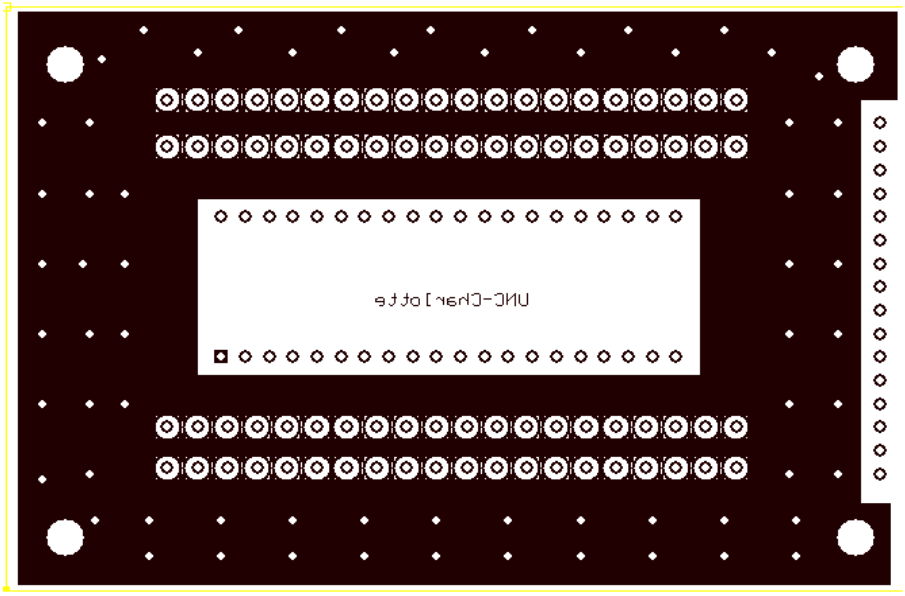

Figure D.2 Back side of printed circuit board used for test setup in Fig. 5.2 and Fig. 5.10

APPENDIX E: BSIM3V3 TRANSISTOR MODEL

T2AK SPICE BSIM3 VERSION 3.1 PARAMETERS

SPICE 3f5 Level 8, Star-HSPICE Level 49, UTMOST Level 8

```
* DATE: Dec 16/02
* LOT: T2AK                WAF: 3202
* Temperature_parameters=Default
.MODEL CMOSN NMOS (                                LEVEL = 49
+VERSION = 3.1          TNOM = 27          TOX = 1.43E-8
+XJ = 1.5E-7            NCH  = 1.7E17      VTH0 = 0.6458545
+K1 = 0.8758326         K2 = -0.0863735    K3 = 19.517147
+K3B = -7.8393631       W0 = 1E-8          NLX = 1E-9
+DVT0W = 0              DVT1W = 0          DVT2W = 0
+DVT0 = 2.7580784       DVT1 = 0.4586878   DVT2 = -0.1580747
+U0 = 455.2375561       UA = 1E-13         UB = 1.826716E-18
+UC = 2.175854E-11      VSAT = 1.525909E5  A0 = 0.6562464
+AGS = 0.1400853        B0 = 2.590299E-6   B1 = 5E-6
+KETA = -4.210703E-3    A1 = 1.408544E-5   A2 = 0.3847639
+RDSW = 1.538425E3      PRWG = 5.265651E-3 PRWB = 0.0326722
+WR = 1                 WINT = 2.938526E-7 LINT = 2.262072E-8
+XL = 0                 XW = 0             DWG = -2.293854E-8
+DWB = 4.66445E-8       VOFF = -0.0103394  NFACTOR =1.0621241
+CIT = 0                CDSC = 2.4E-4      CDSCD = 0
+CDSCB = 0              ETA0 = 0.015919    ETAB = -1.83082E-3
+DSUB = 0.2305335       PCLM = 2.502833    PDIBLC1=1.260663E-3
+PDIBLC2 = 2.660119E-3  PDIBLCB = -0.0238723 DROUT = 5171882
+PSCBE1 = 6.172273E8    PSCBE2 = 1.615457E-4 PVAG = 2.63876E-3
+DELTA = 0.01           RSH = 83.8         MOBMOD = 1
+PRT = 0                UTE = -1.5         KT1 = -0.11
+KT1L = 0               KT2 = 0.022        UA1 = 4.31E-9
+UB1 = -7.61E-18        UC1 = -5.6E-11     AT = 3.3E4
+WL = 0                 WLN = 1            WW = 0
+WWN = 1                WWL = 0            LL = 0
+LLN = 1                LW = 0             LWN = 1
+LWL = 0                CAPMOD = 2         XPART = 0.5
+CGDO = 2.12E-10        CGSO = 2.12E-10    CGBO = 1E-9
+CJ = 4.239616E-4       PB = 0.9873011     MJ = 0.4472349
+CJSW = 3.319142E-10    PBSW = 0.1         MJSW = 0.11571
+CJSWG = 1.64E-10       PBSWG = 0.1        MJSWG = 0.11571
+CF = 0                 PVTH0 = 0.0757071  PRDSW = 141.3097407
+PK2 = -0.0281147       WKETA = -0.0162998 LKETA=5.366192E-4)
*
MODEL CMOSP PMOS                                   (LEVEL = 49
+VERSION = 3.1          TNOM = 27          TOX = 1.43E-8
+XJ = 1.5E-7            NCH = 1.7E17       VTH0 = -0.9655433
+K1 = 0.5357026         K2 = 0.0123939     K3 = 6.143499
```

```
+K3B = -0.89781          W0 = 1E-8              NLX = 4.721735E-8
+DVT0W = 0               DVT1W = 0              DVT2W = 0
+DVT0 = 1.871415         DVT1 = 0.5346187       DVT2 = -0.1378656
+U0 = 226.8100934        UA = 3.369805E-9       UB = 2.060368E-21
+UC = -5.52072E-11       VSAT = 1.747657E5      A0 = 0.9045319
+AGS = 0.1607865         B0 = 1.175028E-6       B1 = 5E-6
+KETA = -1.984718E-3     A1 = 0                 A2 = 0.3
+RDSW = 3E3              PRWG = -0.0326554      PRWB = -9.35987E-3
+WR = 1                  WINT = 3.201584E-7     LINT = 3.818055E-8
+XL = 0                  XW = 0                 DWG = -2.817465E-8
+DWB = 2.003866E-8       VOFF = -0.0845828      NFACTOR=0.7962591
+CIT = 0                 CDSC = 2.4E-4          CDSCD = 0
+CDSCB = 0               ETA0 = 0.3419057       ETAB = -0.1086238
+DSUB = 1                PCLM = 2.2447279       PDIBLC1 = 0.0410495
+PDIBLC2 = 3.907213E-3   PDIBLCB = -0.0396913   DROUT = 0.2037278
+PSCBE1 = 6.981994E9     PSCBE2 = 6.828924E-10  PVAG = 0.0147169
+DELTA = 0.01            RSH = 103              MOBMOD = 1
+PRT = 0                 UTE = -1.5             KT1 = -0.11
+KT1L = 0                KT2 = 0.022            UA1 = 4.31E-9
+UB1 = -7.61E-18         UC1 = -5.6E-11         AT = 3.3E4
+WL = 0                  WLN = 1                WW = 0
+WWN = 1                 WWL = 0                LL = 0
+LLN1 = 1                LW = 0                 LWN = 1
+LWL = 0                 CAPMOD = 2             XPART = 0.5
+CGDO = 2.88E-10         CGSO = 2.88E-10        CGBO = 1E-9
+CJ = 7.247257E-4        PB = 0.9542821         MJ = 0.4961393
+CJSW = 2.703165E-10     PBSW = 0.99            MJSW = 0.2845732
+CJSWG = 6.4E-11         PBSWG = 0.99           MJSWG = 0.2845732
+CF = 0                  PVTH0 = 5.98016E-3     PRDSW = 14.8598424
+PK2 = 3.73981E-3        WKETA = 5.663201E-3    LKETA=-6.872318E-3)
*
```

APPENDIX F: ADS SCHEMATIC USED FOR SIMULATION OF TWO-STAGE AMPLIFIER

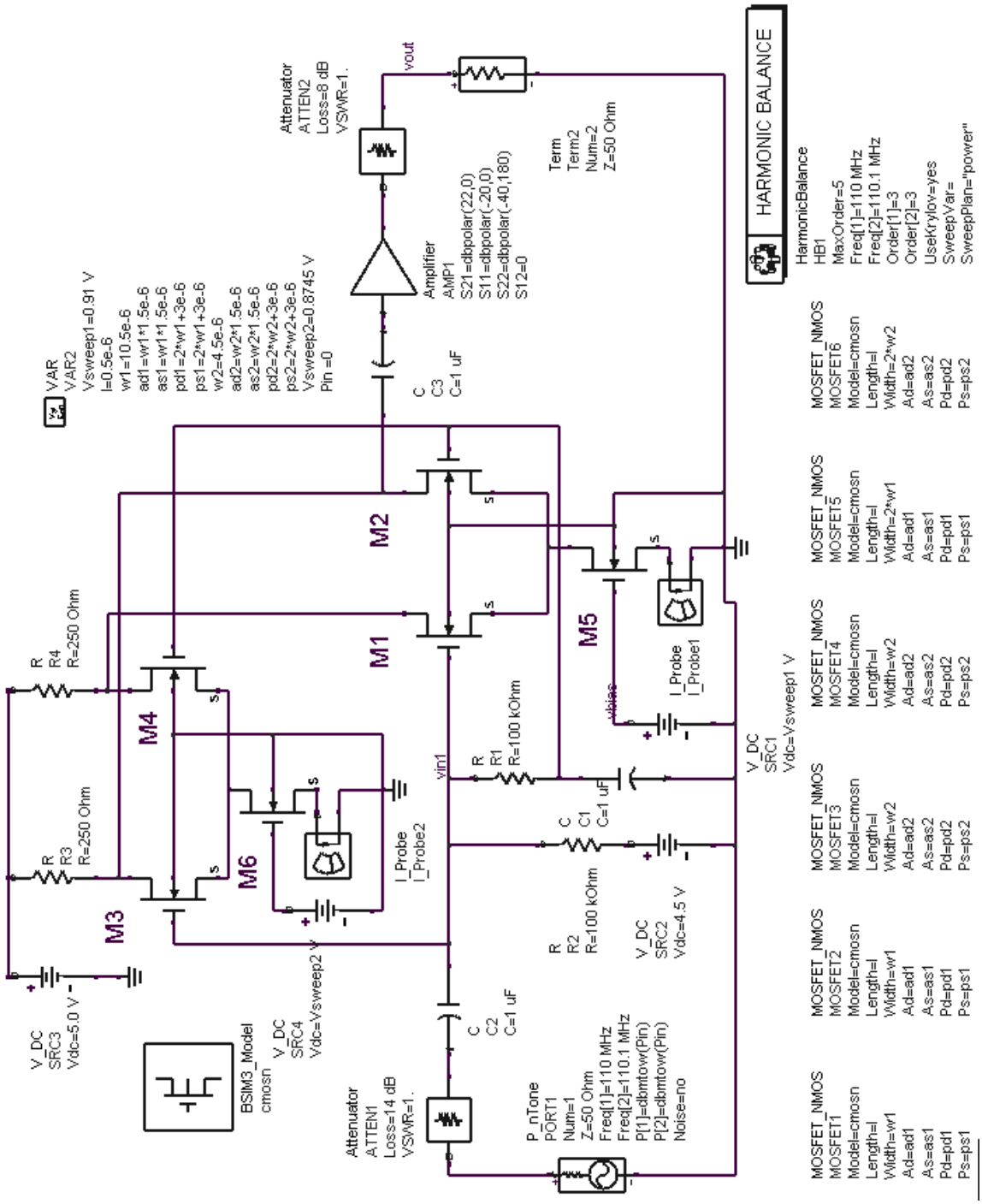

Figure F.1 ADS schematic of cross-coupled differential pair used for simulations with ADS

APPENDIX G: MATHCAD PROGRAM FOR FOUR-STAGE AMPLIFIER LINEARIZATION

ip -4amp7.mcd

Mathcad program to calculate suppression of third order distortion
and gain compression curve for a four-stage amplifier linearization

Forrnulae

p1=linear siganl
p3=3rd order level
G=gain dB, OIP=3rd order output intercept pont in dBm

General compressed gain output power level formula:

$$p1(Pin, G, OIP) := \min(Pin + G, OIP - 12)$$

p1=Output power including compression

General third order intermodulation product output power level formula:

$$p3(Pin, G, OIP) := OIP - 3(OIP - p1(Pin, G, OIP))$$

p3=Third order output distortion

Derived / calculated parameters

$$calcG2(g1) := g1 - 10$$

Rule of thumb for gain G2 of amplifier A2

$$calcAtt(g1, g2) := g1 - g2$$

Attenuator loss

$$calcOip2(OIP1, att, deltaoip) := OIP1 - att - deltaoip$$

Derived value of OIP3 to assure cancellaton of linear signal at output of first subtractor

deltaoip=Difference in output intercept point, derated to inputs of first subtractor

Calculation of third order distortion level at output of amplifier A3

calcP3_3out = Third order distortion level at output of amplifier A3

$$calcP3\_3out(pin, g1, o1, g3, o3) := 20\log\left(10^{\frac{p1(p3(pin, g1, o1), g3, o3)}{20}} + 10^{\frac{p3(p1(pin, g1, o1), g3, o3)}{20}}\right)$$

Calculation of third order distortion level at output of first subtractor (error signal)

calcP3_Errout = Third order distortion level at output of first subtractor (error signal)

$$\text{calcP3\_Errout}(pin, g1, o1, g2, o2, att) := 10 \cdot \log\left[\left[10^{\frac{p3(pin, g2, o2)}{20}} - 10^{\frac{(p3(pin, g1, o1) - att)}{20}}\right]^2\right]$$

Subtraction takes place at voltage level

Calculation of third order distortion level at output of amplifier A4

calcP3_4out = Third order distortion level at output of amplifier A4

$$\text{calcP3\_4out}(pin, g1, o1, g2, o2, att, g4, o4) := p1(\text{calcP3\_Errout}(pin, g1, o1, g2, o2, att), g4, o4)$$

Calculation of third order distortion level of final output

calcFinalP3_out = Third order distortion level at final output

$$\text{calcFinalP3\_out}(pin, g1, o1, g2, o2, att, g3, o3, g4, o4) := 10 \cdot \log\left[\left[10^{\frac{\text{calcP3\_3out}(pin, g1, o1, g3, o3)}{20}} \ldots + \left(-10^{\frac{\text{calcP3\_4out}(pin, g1, o1, g2, o2, att, g4, o4)}{20}}\right)\right]^2\right]$$

Subtracton takes place at the voltage level

Calculation of suppression of third order distortion at final output

suppression = Third order distortion level suppression at final output

$$\text{suppression}(pin, g1, o1, g2, o2, att, g3, o3, g4, o4) := \text{calcP3\_3out}(pin, g1, o1, g3, o3) \ldots + (-\text{calcFinalP3\_out}(pin, g1, o1, g2, o2, att, g3, o3, g4, o4))$$

3rd order distortion of amplifier A3 minus 3rd order IMD at final output

Calculation of linear power at final output

calcPout = linear output power at final output

$$\text{calcPout}(pin, g1, o1, g3, o3) := p1(p1(pin, g1, o1), g3, o3)$$

cascaded amplifier A1 and amplifier A3

Example calculation ==========

Free parameters:

$P := -40$ Input Power

$G1 := 20$ $OIP1 := 10$ Amplifier A1 Parameters

$G3 := 30$ $oip3overdrive := 10$ Amplifier A3 Parameters
oip3overdrive = (OIP1+G3) - OIP3

$deltaoip12 := 10$ deltaoip=Difference in output intercept point derated to inputs of first subtractor

Derived Parameters:

$G2 := calcG2(G1)$ Gain of amplifier A2

$Att := calcAtt(G1, G2)$ Value of necessary attenuation

$OIP2 := calcOip2(OIP1, Att, deltaoip12)$ $OIP3_2$ of amplifier A2

$OIP3 := OIP1 + G3 - oip3overdrive$ $OIP3_3$ of amplifier A3

$OIP4 := OIP3 - 20$ $OIP3_4$ of amplifier A4

$G4 := calcP3\_3out(P, G1, OIP1, G3, OIP3) - calcP3\_Errout(P, G1, OIP1, G2, OIP2, Att)$

Gain of amplifier A4

Summary of parameters of amplifiers A1 to A4

$G1 = 20$	$G2 = 10$	$G3 = 30$	$G4 = 41.743$	$Att = 10$
$OIP1 = 10$	$OIP2 = -10$	$OIP3 = 30$	$OIP4 = 10$	

$calcP3\_3out(P, G1, OIP1, G3, OIP3) = -29.172$ Third order intermodulation distortion IMD3 at output of A3

$calcP3\_Errout(P, G1, OIP1, G2, OIP2, Att) = -70.915$ IMD3 of output first subtractor

$calcP3\_4out(P, G1, OIP1, G2, OIP2, Att, G4, OIP4) = -29.172$ IMD3 of output of amplifier A4

$calcFinalP3\_out(P, G1, OIP1, G2, OIP2, Att, G3, OIP3, G4, OIP4) = -337.154$ IMD3 of final output

$suppression(P, G1, OIP1, G2, OIP2, Att, G3, OIP3, G4, OIP4) = 307.981$ Third order distortion suppression

$calcPout(P, G1, OIP1, G3, OIP3) = 10$ Linear output power at final output

Calculation for input power sweep

Vector Arrays for plots:

$N := 0 .. 40$

$Pin_N := 0 - N$ Input power swept from -40 to 0 dBm

Third order suppression as a function of input power

$Sup_N := suppression(Pin_N, G1, OIP1, G2, OIP2, Att, G3, OIP3, G4, OIP4)$

Linear output power of amplifier A3 as a function of input power

$Pout_N := calcPout(Pin_N, G1, OIP1, G3, OIP3)$

Third order distortion of output of amplifier A3 as a function of input powe

$p33o_N := calcP3\_3out(Pin_N, G1, OIP1, G3, OIP3)$

Third order distortion of output of amplifier A4 as a function of input powe

$p34o_N := calcP3\_4out(Pin_N, G1, OIP1, G2, OIP2, Att, G4, OIP4)$

Third order distortion of final output as a function of input power

$p3o_N := calcFinalP3\_out(Pin_N, G1, OIP1, G2, OIP2, Att, G3, OIP3, G4, OIP4)$

Error signal as a function of input power

$p3error_N := calcP3\_Errou(Pin_N, G1, OIP1, G2, OIP2, Att)$

Final linear output power as a function of input power

$finalP_N := calcPout(Pin_N, G1, OIP1, G3, OIP3)$

APPENDIX H: RESULTS OF PROTOTYPE USING DISCRETE DEVICES

The new concept for linearization was applied to a second prototype and was built with discrete devices in order to gain information about the feasibility of the condition for cancellation described with expression (4.7). A schematic of the prototype is shown in Fig. H.1. The input signal is input to a 180 degree hybrid of type AMT-2 on the left side in Fig. H.1. The hybrid distributes the input signal equally between more linear amplifier A1 of type ERA5-SM at the top of Fig. H.1 and less linear amplifier A2 of type MAR1-SM at the bottom of Fig. H.1. The output signal of A1 is fed into a second 180 degree hybrid of type AMT-2 on the right side of Fig. H.1 and output signal of A2 is attenuated by a 12 dB pi-attenuator and then input to the 180 degree phase shift port of the second hybrid. Output signal of A2 experiences a phase shift of 180 degree passing through the second hybrid. As a result, the final output signal is the difference between output signals of A1 and A2. The second hybrid at the output acts as subtractor.

All devices used are products of Minicircuits [36] and mounted on a printed circuit board. A layout of the top view of the PCB is depicted in Fig. H.2. The transmission lines in the layout were designed as coplanar wave guides with 50 Ω impedance. The back side of the PCB is completely metalized and acts as ground plane. The material of the PCB is FR4 with relative dielectric constant of 4 to 5. The PCB is equipped with SMA connectors for interconnection with other devices.

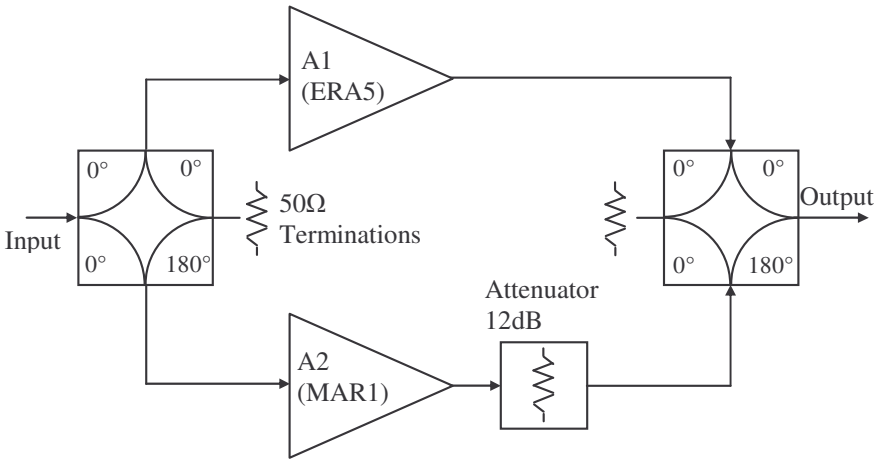

Figure H.1 Block diagram of new linearization method using discrete devices: input signal is split equally in-phase with a 180 degree hybrid and fed to amplifiers A1 and A2 and output signals are subtracted with hybrid. The fourth port of hybrids (AMT-2) is terminated with 50Ω resistors. The 12dB attenuator has pi form and is passive. Amplifiers A1 and A2 are ERA5-SM and MAR1-SM.

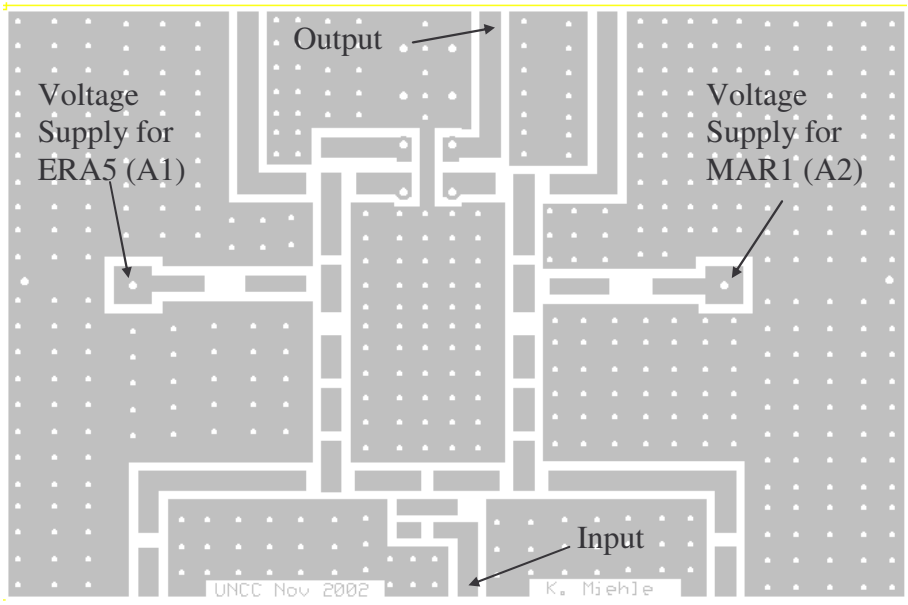

Figure H.2 Top layout for test board for discrete devices used in test setup in Fig. H.3; traces are coplanar 50Ω transmission lines.

Supply voltage and current of amplifiers A1 and A2 were adjusted to meet the third order distortion cancellation condition described with expression (4.7). The pi-attenuator also was added to reach cancellation condition. A summary of supply voltage and currents is shown in Table H.1.

Table H.1 Summary of supply voltages and currents of amplifiers A1 and A2 from Fig. H.1.

	Amplifier A1 (ERA5)	Amplifier A2 (MAR1)
Supply Voltage	7.8 V	7 V
Supply Current	30 mA	19 mA

The test circuit on the PCB was tested utilizing a two-tone test. The test setup is shown in Fig. H.3. Two frequencies were set to 125 MHz and 125.1 MHz with the two signal generators of type HP E4400B on the left side in Fig. H.3. Each generator has a 10 dB attenuator at the output to isolate the two generators. The purpose of the following 6 dB attenuator is to improve impedance matching to the test circuit. Further on, the output of the 6 dB attenuator is connected to the input of the test circuit consisting of the circuit shown in Fig. H.1. Finally, the measured results were displayed with a spectrum analyzer of type HP E4401B shown on the right side in Fig. H.3. Furthermore, the two signals at frequencies 125 MHz and 125.1 MHz, as input of the two-tone test setup, were adjusted to a power level of –21 dBm at the input of the test circuit.

After adjusting the amplifiers and the attenuator, the measured results shown in Figs H.4 to H.6 were achieved. First, amplifier A2 (MAR1) was turned off and only the frequency spectrum of amplifier A1 (ERA5) was measured. This frequency spectrum is

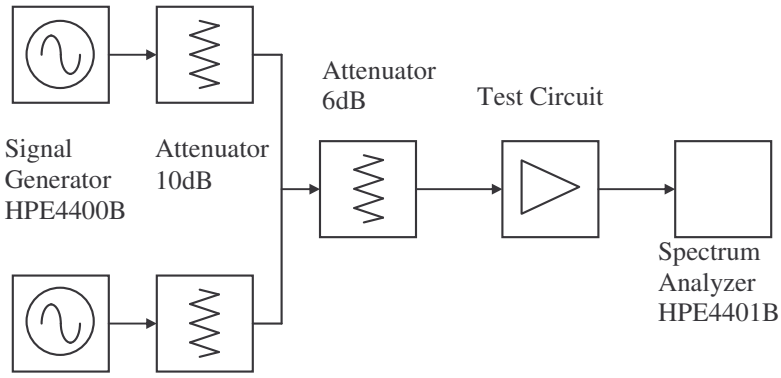

Figure H.3 Test setup for circuit in Fig. H.1 using discrete devices: the two-tone test signal is created with the generators of type HP E4400B on the left side, two 10dB attenuators are passive and isolate the signal generators; the 6dB attenuator is a passive device to match to the test circuit. The output of the test circuit is connected the spectrum analyzer of type HP E4401B on the right side.

depicted in Fig. H.4. The spectrum shows the two fundamental tones in the center with 125 MHz and 125.1MHz and a power level of -11.5 dBm. Due to nonlinearities in the amplifier, third order intermodulation products appear at frequencies 124.9 MHz and 125.2 MHz (outermost spectral lines in Fig. H.4) with power levels of –60.6 dBm.

Secondly, amplifier A1 was deactivated and only amplifier A2 was measured. Fig. H.5 illustrates the frequency spectrum of amplifier A2 only. Similar to Fig. H.4, the innermost spectral lines are the fundamental tones at frequencies 125 MHz and 125.1 MHz with power levels of –20.4 dBm. The two outermost spectral lines in Fig. H.5 are third order intermodulation products, created as a result of third order nonlinearities in amplifier A2. Again, the IMD3 products appear at 124.9 MHz and 125.2 MHz and have power levels of –61.8 dBm.

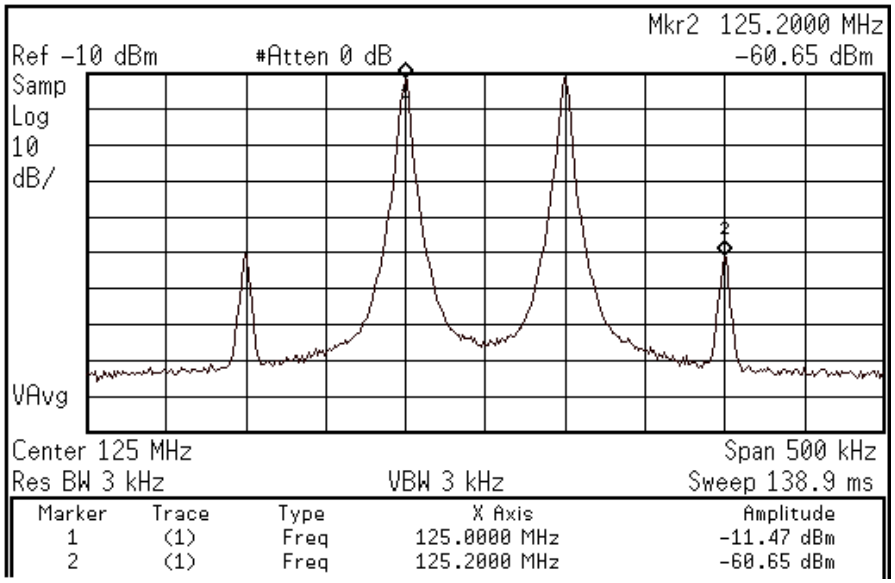

Figure H.4 Measured output frequency spectrum of amplifier A1 (ERA5) at 125MHz using test setup from Fig. H.3: amplifier A2 was deactivated. Innermost lines are fundamental tones, outermost lines are third order distortion products.

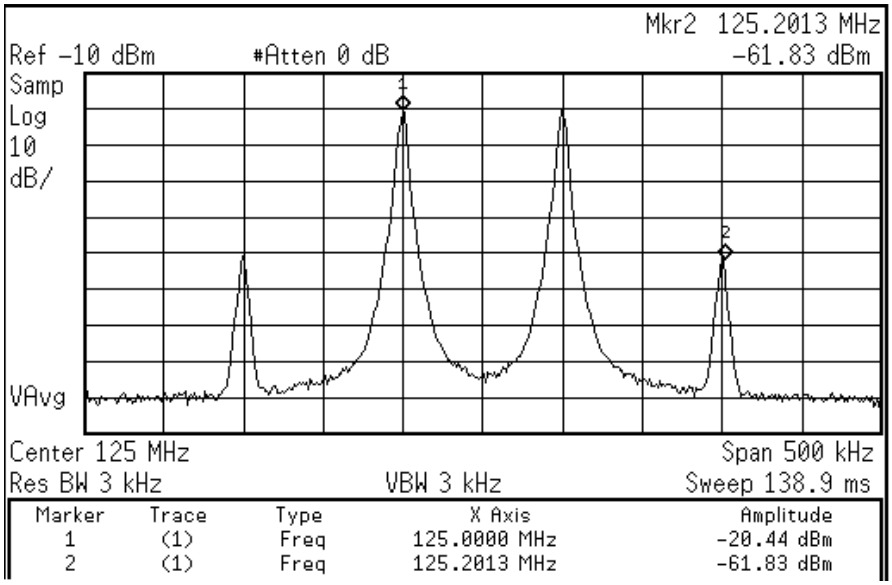

Figure H.5 Measured output frequency spectrum of amplifier A2 (MAR1) at 125MHz using test circuit of Fig. H.3: amplifier A1 was deactivated. Innermost lines are fundamental components, outermost lines are third order distortion products.

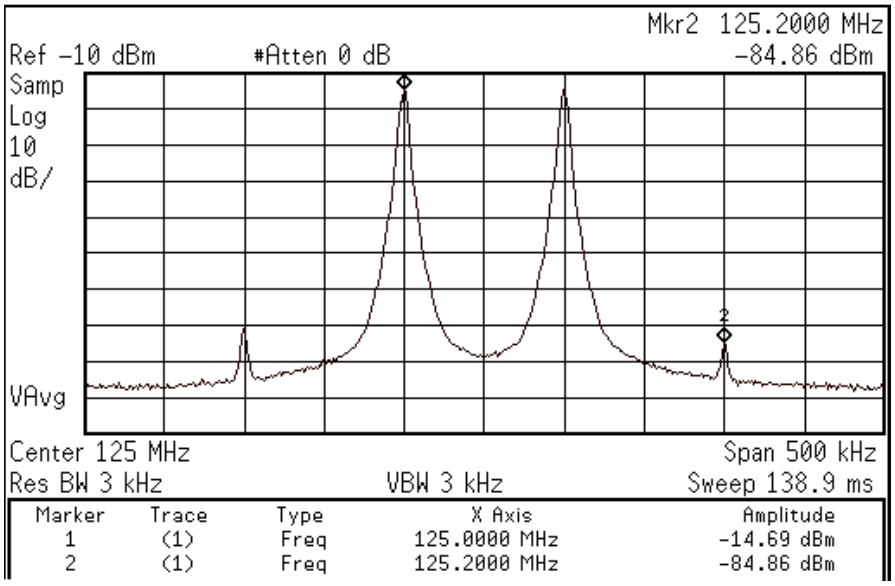

Figure H.6 Measured output frequency spectrum at 125MHz of amplifiers A1+A2 using test setup of Fig. H.3: innermost lines are fundamental components and outermost lines are third order distortion

Comparing the frequency spectra in Fig. H.4 and Fig. H.5, one can see that the IMD3 products are almost equal (-60.6 dBm and –61.8 dBm). By subtracting the output signals of A1 and A2 via a 180 degree hybrid, reduction of third order distortion is achieved. The resulting frequency spectrum is depicted in Fig. H.6. In Fig. H.6, the two spectral lines in the center are the fundamental tones with power levels of –14.7 dBm and the spectral lines on both sides are third order intermodulation products. The frequencies of the spectral lines did not change compared to the previous measurements. The third order distortion products are reduced to –84.8 dBm. Comparing this power level to the power level of third order distortion of amplifier A1 in Fig. I.4, a suppression of 24.2 dB was reached, corresponding to an improvement in OIP3 of approximately 7.4 dB.

Comparing the fundamental tones of Fig. H.4 and Fig. H.6, one can also see that subtraction results also in loss of output power as expected. Not only third order

distortion products are subtracted but also the fundamental components, as result, a loss in output power of approximately 3 dB was measured. Table H.2 gives an overview of all measured results from Figs H.4 to H.6 for the test circuit in Fig. H.1 using the test setup in Fig. H.3.

Table H.2 Summary of measured results for 125 MHz using circuit in Fig. H.1 with test setup in Fig. H.3

	P_{in} (dBm) (at test circuit)	P_{out} (dBm)	Gain (dB)	P_{3rd} (dBm)	OIP3 (dBm)
Amplifier A1 (ERA5, Fig. H.4)	-21	-11.5	9.5	-60.6	13.0
Amplifier A2 (MAR1, Fig. H.5)	-21	-20.4	1.4	61.8	0.3
Amplifier A1+A2 (Fig. H.6)	-21	-14.7	6.3	-84,8	20.4
Comparison: A1 (Fig. H.4) with A1+A2 (Fig. H.6)			Loss in gain: 3.2 dB	Suppression: 24.2 dB	Improvement in OIP3: 7.4 dB

In Table H.2, P_{in} in the first column is the power level of the two tones at frequencies 125 MHz and 125.1 MHz at the input of the test circuit. The linear output power of the fundamental components in Figs. H.4 to H.6 (two innermost spectral lines) is denoted P_{out} and is shown in the second column in Table H.2. Gain, calculated as P_{out}-P_{in}, is listed in the third column in Table H.2. Next to the gain, in the fourth column from the left, is the power level of third order distortion listed and denoted as P_{3rd}. In Figs. H4 to H6, P_{3rd} is the power level of the outermost spectral lines. Finally, in the column on the

right side in Table H.2 is the third order output intercept point OIP3 listed. OIP3 was calculated with equation (3.28) using measured results from Table H.2.

Furthermore, the first row in Table H.2 shows measured results of amplifier A1 only. Amplifier A2 was deactivated. Measured results of amplifier A2 are listed in the second row in Table. H.2. Amplifier A1 was deactivated during these measurements. The third row summarizes measured results when both amplifiers were turned on for cancellation of third order distortion. The supply voltage and currents for this situation are listed in Table H.1. The row at the bottom of Table H.2 gives results of a comparison of measured and calculated results of amplifier A1 in the first row and results of the third row when both amplifiers are turned on for cancellation.

To prove the cancellation condition, measured results were substituted into equation (4.7). For review, (4.7) is $2(OIP2_1\text{-}OIP3_2)=3(G_1\text{-}G_2)$. Substituting OIP3 of both amplifiers into the left side of (4.7) results in

$2(OIP2_1\text{-}OIP3_2)=2(13\text{-}0.3)=25.4$ dB

Substituting measured results into the right side of (4.7) gives

$3(G_1\text{-}G_2)=3(9.5\text{-}1.4)=24.3$ dB

The results show that equation (4.7) holds well for the tested circuit. Note, that the devices are off the shelf and are not particularly designed to fulfill the cancellation condition under the above described scenario, but the results indicate that the condition of cancellation described with (4.7) is very general.