A Base Control Doherty Power Amplifier Design
for Improved Efficiency in GSM Handsets

by
Darren W. Ferwalt

A THESIS
submitted to
Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented December 10, 2003
Commencement June 2004
© Copyright by Darren W. Ferwalt
December 10, 2003
All Rights Reserved
AN ABSTRACT OF THE THESIS OF

Darren W. Ferwalt for the degree of Master of Science in Electrical and Computer Engineering presented on December 10, 2003.

Title: A Base Control Doherty Power Amplifier Design for Improved Efficiency in GSM Handsets.

Abstract approved: ____________________________________________

Andreas Weisshaar

In cellular telephones, RF power amplifiers consume a significant part of the total phone current. Reducing the average PA current will extend battery life. In GSM systems the handset is commanded by the base station to transmit at power levels lower than full power much of the time. A Doherty amplifier can be used to improve efficiency at these lower power levels, providing a lower average PA current when evaluated over the entire power range.

A Doherty amplifier is presented that uses the DC base voltage to control the output power. The base control amplifier is first studied separately, then two such amplifiers are combined to create a Doherty amplifier. An analysis of battery life while the phone is on a call, or talk time, is then performed. Talk time for the base control Doherty amplifier is compared to an ideal class B amplifier and to a Doherty amplifier composed of two class B amplifiers. Finally a base control Doherty amplifier is implemented at 30 MHz, both in simulation and hardware.
ACKNOWLEDGMENT

I would like to thank Tom Apel who introduced me to this subject material and suggested a GSM Doherty amplifier as a thesis topic. His help at various times throughout the research was invaluable.

I am most grateful to TriQuint Semiconductor and Dennis Criss for giving me the time and support to finish this task. Thanks also to Chris Stephens and Sepideh Rezania at TriQuint Semiconductor, who helped educate me on the finer points of cell phone talk time analysis.

To my committee, I am grateful for all of the time they spent from their busy schedules. Thanks especially to Dr. Andreas Weisshaar for supporting my program and providing “World Olympics” quality pep-talks, and to Dr. Fred Weiss, who has been a true mentor in the finest of traditions. Thanks also to Dr. Rick Campbell for his expertise, advice and encouragement along the way.

And of course to my wife, Mary Kay, I am most grateful. She not only kept me going through this process, but also helped with proofreading and formatting on several occasions. Thanks also to Corinn, Nolan, and Ethan, the finest and most patient gaggle of kids a guy could ask for.
# TABLE OF CONTENTS

1. **INTRODUCTION** ..............................................................................................................1  
   1.1. Motivation ..............................................................................................................1  
   1.2. A Proposed Solution ..........................................................................................3  
   1.3. Power Amplifier Classes ..................................................................................6  

2. **LITERATURE REVIEW** .................................................................................................9  
   2.1. The Doherty Amplifier ....................................................................................9  
   2.2. Recent Doherty Amplifier Literature ...........................................................15  
   2.3. Other Low Power Efficiency Enhancement Techniques ..............................19  

3. **CIRCUIT ANALYSIS** ...................................................................................................21  
   3.1. Problem Statement ..........................................................................................21  
   3.2. Optimizing Load Impedance for Efficiency ..................................................25  
   3.3. Doherty Load Modulation .............................................................................29  
   3.4. The Base Control Amplifier ..........................................................................33  
   3.5. Applying Base Control to the Doherty Amplifier .........................................42  
   3.6. Talk Time Analysis .......................................................................................46  

4. **PROTOTYPE DESIGN** ..................................................................................................49  
   4.1. Topology and Device Technology ....................................................................49  
   4.2. Bias Circuits ....................................................................................................53  
   4.3. RF Core Designs ..............................................................................................56  
   4.4. Impedance Inverters .......................................................................................58  
   4.5. Simulation Results .........................................................................................60  

5. **PROTOTYPE RESULTS** ................................................................................................66  
   5.1. Hardware Construction ...................................................................................66  
   5.2. Hardware Results ............................................................................................68  
   5.3. Lessons Learned ...............................................................................................71
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. CONCLUSION</td>
<td>72</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>73</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.1</td>
<td>CDMA handset PA power statistics in an urban environment</td>
</tr>
<tr>
<td>1.2</td>
<td>Ideal class B efficiency vs. power out</td>
</tr>
<tr>
<td>1.3</td>
<td>Doherty amplifier</td>
</tr>
<tr>
<td>1.4</td>
<td>Base controlled power control loop</td>
</tr>
<tr>
<td>1.5</td>
<td>Ideal class A, B, or C amplifier circuit</td>
</tr>
<tr>
<td>1.6</td>
<td>Conduction angles for a class A, B, or C amplifier</td>
</tr>
<tr>
<td>2.1</td>
<td>Active load-pull concept</td>
</tr>
<tr>
<td>2.2</td>
<td>Doherty action</td>
</tr>
<tr>
<td>2.3</td>
<td>Output voltage and current relationships from Doherty’s analysis</td>
</tr>
<tr>
<td>2.4</td>
<td>Doherty’s efficiency results</td>
</tr>
<tr>
<td>2.5</td>
<td>Ideal Doherty efficiency vs. power out</td>
</tr>
<tr>
<td>2.6</td>
<td>Practical linear Doherty efficiency</td>
</tr>
<tr>
<td>3.1</td>
<td>Urban PDF for GSM PCS1900 band</td>
</tr>
<tr>
<td>3.2</td>
<td>Ideal Class B collector current (DC), scaled to 50% efficiency at 33 dBm</td>
</tr>
<tr>
<td>3.3</td>
<td>Class B waveforms</td>
</tr>
<tr>
<td>3.4</td>
<td>Optimum class B load impedance vs. power out, Vcc = 3.5V</td>
</tr>
<tr>
<td>3.5</td>
<td>Reduced output power waveforms and power dissipation</td>
</tr>
<tr>
<td>3.6</td>
<td>Output voltage and current relationships from Doherty’s analysis, α = 0.25</td>
</tr>
<tr>
<td>3.7</td>
<td>Doherty action</td>
</tr>
<tr>
<td>3.8</td>
<td>Load presented to Q1 in Doherty amplifier, Vcc = 3.5V, α = 0.25</td>
</tr>
<tr>
<td>3.9</td>
<td>Proposed base control circuit and its equivalent circuit for analysis</td>
</tr>
</tbody>
</table>
**LIST OF FIGURES (Continued)**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.10 Base current and voltage waveforms as $V_{PC}$ is reduced</td>
<td>38</td>
</tr>
<tr>
<td>3.11 Base control collector current waveforms</td>
<td>39</td>
</tr>
<tr>
<td>3.12 Base control efficiency vs. power out</td>
<td>42</td>
</tr>
<tr>
<td>3.13 Base control Doherty efficiency</td>
<td>45</td>
</tr>
<tr>
<td>4.1 Basic HBT cell structure</td>
<td>51</td>
</tr>
<tr>
<td>4.2 Prototype circuit topology</td>
<td>53</td>
</tr>
<tr>
<td>4.3 Doherty fundamental current relationships for $\alpha = 0.25$</td>
<td>54</td>
</tr>
<tr>
<td>4.4 Simulated fundamental currents vs. $V_L$ and $V_{PC}$</td>
<td>56</td>
</tr>
<tr>
<td>4.5 Final prototype simulation circuit</td>
<td>60</td>
</tr>
<tr>
<td>4.6 Ideal and simulated output stage ($Q_{2LP} + Q_{2HP}$) efficiency vs. power out</td>
<td>61</td>
</tr>
<tr>
<td>4.7 Simulated efficiency of all stages of the base control Doherty</td>
<td>62</td>
</tr>
<tr>
<td>4.8 Doherty action: constant voltage saturation in transition region</td>
<td>63</td>
</tr>
<tr>
<td>4.9 Base control PA efficiency, simulation and ideal</td>
<td>64</td>
</tr>
<tr>
<td>4.10 $Q_{2HP}$ base voltage and current waveforms</td>
<td>64</td>
</tr>
<tr>
<td>4.11 Fundamental magnitude of $Q_{1HP}$ and $Q_{2HP}$ collector currents</td>
<td>65</td>
</tr>
<tr>
<td>5.1 Final hardware</td>
<td>67</td>
</tr>
<tr>
<td>5.2 Close-up of transistors</td>
<td>67</td>
</tr>
<tr>
<td>5.3 Hardware efficiency results</td>
<td>69</td>
</tr>
<tr>
<td>5.4 Collector voltage waveforms measured with oscilloscope</td>
<td>70</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.1</td>
<td>Summary of class A, B, and C amplifiers</td>
</tr>
<tr>
<td>3.1</td>
<td>Summation from equation (15)</td>
</tr>
<tr>
<td>3.2</td>
<td>Improvement in average PA current</td>
</tr>
<tr>
<td>3.3</td>
<td>Summary of average PA current and talk time improvement</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

1.1. Motivation

The “talk time” of a cell phone, defined as the time it takes to discharge the battery while on a call, is an important metric to the success of a phone. The power amplifier (PA) consumes more current than any other component, so by maximizing the PA’s efficiency over all of its operating conditions, talk time can be improved.

Cellular phone systems require the PA to put out a range of power levels, depending on distance to the base station and path conditions. Typical usage patterns can be measured, giving the probability that a PA will be required to transmit at a given power level. For example, Fig. 1.1 shows IS-95 CDMA\(^1\) handset power level

![Probability Density](image)

Fig. 1.1 CDMA handset PA power statistics in an urban environment [1].

\(^1\) Code Division Multiple Access, IS-95 is the US standard for the original CDMA cellular system.
statistics in an urban environment [1]. GSM\textsuperscript{2} power levels range from 0 to 33 dBm, and a new study of GSM power level statistics, discussed in Chapter 3, shows a similar distribution centered around 16 dBm. Therefore, to maximize battery life efficiency should be high throughout the power range.

GSM PA manufacturers have reached a practical maximum of 50-60\% power added efficiency (PAE\textsuperscript{3}) at full power. However, lower power efficiency has been largely ignored, possibly due to the lack of published information about GSM power statistics. Fig. 1.2 shows the theoretical efficiency vs. power out (Pout) for a single transistor class B power amplifier. The lower line in Fig 1.2 is a scaled version of the class B curve, such that the maximum efficiency is 50\%. This curve closely matches

![Fig. 1.2 Ideal class B efficiency vs. power out.](image)

\textsuperscript{2} Global Specification for Mobile communications, the predominant cellular system in Europe.

\textsuperscript{3} PAE is defined as (RF power out - RF power in) \div DC power supplied.
the measured efficiency of a multi-stage, GSM PCS1900\(^4\) PA. Comparing this lower
curve to a curve similar to Fig. 1.1 centered around 16 dBm suggests that low power
efficiency will play an important role in talk time. Note that the scaled efficiency at
16 dBm is 7%.

1.2. A Proposed Solution

One method to achieve improved PAE at lower power levels was developed by
William Doherty in the 1930s [2]. The Doherty amplifier is composed of two
amplifiers in parallel, where at full power both are on and at low power one is off (Fig.
1.3). The two amplifiers are configured such that as the power out of PA #2 is
reduced, the load impedance of PA #1 is increased, allowing it to operate at higher
efficiency at a lower power. The combined efficiency at reduced power is therefore

\[ L = \frac{\lambda}{4} \]

\[ Z_0 = Z_{INV} \]

\[ R_L \]

Fig. 1.3 Doherty amplifier.

\( ^4 \) PCS1900 (Personal Communication System), 1900 MHz GSM band in the US.
improved over that of a single amplifier. Detailed operation of the Doherty amplifier will be covered in Section 2.2.

Doherty’s original design used amplifiers of equal power capacity, which resulted in a region of high efficiency from full power to 6 dB below full power. This suited Doherty’s application at the time (fixed power AM broadcast stations), with typical envelope characteristics that mostly remained within this 6 dB window. Since then, Doherty amplifiers have been built for many applications including, most recently, high efficiency linear PAs for cellular digital modulation schemes [3]-[14], which require operation over a wide power range. By adjusting the power capacity ratio of the two sides, the Doherty load modulation effect can be extended to more than 12 dB below full power (this is called an extended Doherty amplifier [15], [3]).

The PA described in this thesis is a Doherty amplifier where both amplifiers are operated in saturated mode (“saturation” is the condition when the voltage waveform at the collector swings from near zero volts to two times the battery voltage, the maximum possible in an ideal class A, B or C amplifier). Although the Doherty amplifier is valuable as a way to increase the efficiency of linear5 PAs, there is no published literature about using the technique to improve the efficiency at backed off power levels of constant envelope modulation schemes, such as GSM. By removing the linearity requirement, the two amplifiers in the Doherty configuration can be run in saturated mode and utilize other nonlinear methods to optimize low power efficiency.

5 Linearity here is defined as a linear relationship between RF input and output power.
One such nonlinear technique widely used in GSM handset PAs is that of controlling the output power by adjusting the base bias of one or more amplifier stages while the RF input level is held constant. When used with a power control loop composed of an output coupler and feedback loop (Fig. 1.4), accurate power control can be achieved. Using base power control, the bias point is naturally set to the minimum at each power level, raising the efficiency throughout the power range.

One weakness of the Doherty amplifier is the need for different gain slopes between PA #1 and PA #2 to achieve the optimum load modulation effect (described in Section 2.1). Using base control for each amplifier in the Doherty configuration provides a solution to this problem, using only DC control circuits and a fixed RF input level. The required power level from each amplifier needed to satisfy the Doherty requirements at each output power level can now be achieved by adjusting its base voltage. If necessary, reasonably sophisticated bias functions could be designed.
into a silicon bias controller integrated circuit (IC), whose cost would be small compared to the rest of the PA module.

The prototype design in this thesis reflects, as much as possible, the practical design constraints in a PA module for a GSM phone. Quarter wavelength lines are impractical in a PA module at cellular frequencies, so lumped-element impedance inverters are used instead of quarter wavelength lines, and only lumped element matches are used. Each amplifier is a two-stage design to provide enough gain for a realistic cellular PA application.

The main exception to a practical GSM PA design is choosing 30 MHz as the design frequency. This allows easier analysis of the basic concepts by minimizing the effects of parasitic inductance and capacitance in the circuit. It also allows for easier discrete implementation (fewer distributed effects make it easier to “glue” components together quickly) and the ability to easily examine voltage waveforms at the outputs. Gallium Arsenide (GaAs) Heterojunction Bipolar Transistors (HBTs) are used in the 30 MHz prototype hardware. This should ease the transition to 1900 MHz, since these same transistors are to be used at that frequency.

1.3. Power Amplifier Classes

Before reviewing the appropriate literature, a short review of class A, B, and C power amplifiers is in order. The simplified circuit in Fig. 1.5 meets the two requirements for a class A, B or C amplifier: (1) the input drive current is assumed to be sinusoidal, such that the collector current maintains a sinusoidal shape when the transistor is “on” and conducts no current when it is in cutoff; and (2) all harmonic
currents are assumed to be shorted to ground at the collector, as represented by the ideal parallel resonant circuit, tuned to the RF frequency of operation, and terminated by the ideal RF ground at VCC. After all harmonic currents are terminated, the resulting current (I in Fig. 1.5) is sinusoidal, and creates a sinusoidal voltage waveform as it terminates in the load resistance, RL.

Fig. 1.5 Ideal class A, B, or C amplifier circuit (class B waveforms shown).

The class of a PA is determined by its conduction angle, which is defined as the portion of each RF sine wave the transistor is “on”. For example, a conduction angle of $\gamma = 2\pi$ means that the transistor conducts current from collector to emitter throughout the entire sine wave, i.e., the transistor is continuously “on” (Fig. 1.6). This is defined as class A operation, and requires the bias voltage, $V_{PC}$, to be sufficiently high to maintain forward base current throughout the RF cycle. Class B is defined as having a conduction angle of $\gamma = \pi$, or when the transistor conducts for half of the RF cycle. $V_{PC}$ is reduced for class B operation so that the base-emitter junction remains
forward biased for half of the RF cycle (see waveforms in Fig. 1.5). Class C operation is defined as having any conduction angle less than $\pi$. Operation with a conduction angle between $\pi$ and $2\pi$ is often referred to as class AB.

![Waveforms](image)

Fig. 1.6 Conduction angles for a class A, B, or C amplifier.

The maximum collector efficiencies for ideal class A, B and C amplifiers [16] are shown below in Table 1.1, and will be discussed in more detail in Section 2.2.

<table>
<thead>
<tr>
<th>Amplifier Class</th>
<th>Conduction angle</th>
<th>Ideal Collector Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class A</td>
<td>$\gamma = 2\pi$</td>
<td>50%</td>
</tr>
<tr>
<td>Class AB</td>
<td>$\pi &lt; \gamma &lt; 2\pi$</td>
<td>50% to 78.5%</td>
</tr>
<tr>
<td>Class B</td>
<td>$\gamma = \pi$</td>
<td>78.5%</td>
</tr>
<tr>
<td>Class C</td>
<td>$\gamma &lt; \pi$</td>
<td>78.5% to 100%</td>
</tr>
</tbody>
</table>

Table 1.1 Summary of class A, B, and C amplifiers.
2. LITERATURE REVIEW

2.1. The Doherty Amplifier

The 1930s was an age of 100 kilowatt AM broadcast stations in the rapidly developing world of radio communications. Large tube amplifiers would put out this 100 kilowatts of power with efficiencies near 30%, so hundreds of kilowatts would be dissipated in the devices. Reducing the expense of water cooling systems was one of Doherty’s main objectives in developing his breakthrough technique. He was able to improve plate efficiency from 35% to 65%, reducing the size of the plate supply by half and the water cooling system by a factor of four [2].

Doherty’s technique uses what Steve C. Cripps calls the “active load-pull” concept [16], where current from one transistor (or tube) modifies the load seen by the other device (Fig. 2.1). For example, as RF current from Q2 goes through the load

Fig. 2.1 Active load-pull concept.
resistance, \( R_L \), the voltage across \( R_L \) increases. The load resistance then “seen” by Q1, defined as \( V_L \) divided by \( I_1 \), goes up since a higher voltage is experienced than would be from the current due to Q1 alone. The value of the apparent resistance seen by each transistor is determined by a ratio of these currents through the load [16]:

\[
R_1 = R_L \frac{I_1 + I_2}{I_1} \quad (1)
\]

\[
R_2 = R_L \frac{I_1 + I_2}{I_2} \quad (2)
\]

Here and elsewhere, all variables used in equations are phasor magnitudes unless otherwise noted.

Using this effect, the current from one transistor can be used to manipulate the load resistance seen by the other. This concept is what allows large arrays of parallel transistor cells to work together, each cell supplying a part of the current and “seeing” an apparent resistance proportionately higher than the load resistance.

For a single transistor class B amplifier with a fixed load resistance, maximum efficiency occurs at full power where the RF voltage swing is at its maximum (called saturation). Efficiency falls off rapidly as the power out is reduced because the transistor comes out of saturation. If at lower power the load resistance could be increased, saturation, thus maximum efficiency, would be restored (saturation and the relationship between efficiency and load impedance will be discussed extensively in Section 3.1). But the active load-pull concept discussed above moves the load resistance seen by Q1 the wrong way, that is, lower in value as the power from Q2 is reduced. So Doherty used an impedance inverter network (equivalent to a quarter-
wave transmission line) between the devices to reverse the active load-pull effect, and
provide a rising $R_1$ as $I_2$ is reduced (Fig. 2.2). If this rising $R_1$ is coupled with a
decreasing RF drive to Q1, maximum efficiency is maintained until $I_2$ is finally zero.
Doherty’s technique allows all of these simultaneous effects to happen automatically
as only the RF drive level to both sides is reduced.

The impedance inverter in Fig. 2.2 also introduces a $90^\circ$ phase lag in the output
current from Q1. The RF input signal for Q2 must therefore also lag by $90^\circ$ in order
for $I_2$ and $I_{1T}$ to combine in phase. This phase lag is normally achieved with a second
impedance inverter at the input, as shown in Fig. 1.3. The governing equations for the
impedance inverters are [15]:

$$R_1 R_{1T} = Z_{INV}^2$$  \hspace{1cm} (3)

$$I_1 = \frac{V_L}{Z_{INV}}$$  \hspace{1cm} (4)
\[ I_{1T} = \frac{V_i}{Z_{\text{INV}}}, \quad (5) \]

where \( Z_{\text{INV}} \) is the characteristic impedance of the impedance inverter. Note from (4) and (5) that the voltage on one side of the inverter is proportional to the current on the other side, but the voltages on both sides are not directly related. This property allows \( V_1 \) to remain fixed at saturation while \( V_L \) is varied, an important part of the Doherty action.

Doherty chose both amplifiers to put out equal power at full power, and chose \( Z_{\text{INV}} \) to be twice the load resistance. If \( \alpha \) is defined as the ratio of the power from Q1 to the total power out (at full power), then

\[ Z_{\text{INV}} = 2R_L \quad (6) \]

\[ \alpha = \frac{P_{Q1MAX}}{P_{\text{MAX}}} = \frac{P_{\text{MAX}}/2}{P_{\text{MAX}}} = \frac{1}{2} \quad (7) \]

where \( P_{\text{MAX}} \) is the total power from both sides at full power, and \( P_{Q1MAX} \) is the power from Q1 at full power. \( R_L \) was chosen to provide the appropriate saturated voltage swing at the desired power and applied collector (or plate) voltage, as described in Section 3.2. From the active load-pull concept in (1), at full power the impedance presented at both \( R_{1T} \) and \( R_2 \) will be 2 times \( R_L \), since \( I_{1T} = I_2 \). Since the characteristic impedance of the impedance inverter, \( 2R_L \), matches that of its load (\( = 2R_L \)), no transformation will occur, allowing Q1 to saturate at \( P_{Q1MAX} \) at its maximum efficiency. Q2 is also saturated and at maximum efficiency.
As the current from Q2 is reduced, V₁ is held constant due to the Doherty action and maximum efficiency is maintained in Q1 (though it is not in Q2). When I₂ is finally reduced to zero, R₁ is transformed to its highest value. Using (3) and (6),

\[ R₁(I₂ = 0) = \frac{Z_{inv}^2}{R_L} = \frac{(2R_L)^2}{R_L} = 4R_L. \] (8)

This higher impedance allows Q₁ to saturate at P_{MAX}/4, thus achieving maximum efficiency at 6 dB below full power.

Doherty summarized the voltage and current relationships for the two devices in Fig. 2.3. Note that tube 1 maintains constant RF plate voltage after the transition point, which is defined as the point where tube 2 first begins to conduct current. Both RF current curves are linear functions, but since tube 2 turns on much later than tube 1 and ends up at the same final value, the slopes are different. Assuming ideal transconductance devices, the output currents shown in Fig. 2.3b dictate the RF voltages that must be applied to the device inputs. As discussed earlier, this is a key

![Fig. 2.3 Output voltage and current relationships from Doherty’s analysis [2.](image)]
weakness of the Doherty amplifier, the need for different input drive slopes for each device. Doherty solved this problem by using a class C peaking amplifier, which has a higher gain slope than his class B carrier amplifier.

In the transition region, defined as the region where $I_2$ is ramping up, the overall efficiency is a composite of the two amplifiers’ efficiencies. Since the peaking amplifier does not have the benefit of load modulation, its efficiency falls off normally as it comes out of saturation. The composite efficiency then falls off below full power, as shown in Fig. 2.4. It is important to note that more recent research (Fredrick Raab’s detailed analysis of the Doherty amplifier in 1987 [15], discussed in the next section) predicts a second “peak” in efficiency at the transition point, since the peaking amplifier is completely off and the carrier amplifier is still at maximum efficiency. A “twin peaks” shape results [16], with a “dip” in between the peaks.

![Fig. 2.4 Doherty’s efficiency results](image)
(shown in the next section). Doherty’s efficiency curve (Fig. 2.4) shows no dip at all, possibly due to tweaking the design on the bench for best overall linear performance, therefore not fully satisfying the relationship of Fig 2.3b.

2.2. Recent Doherty Amplifier Literature

Much has been published about applying Doherty’s concept to wireless and microwave frequencies. Recent linear wireless standards such as CDMA, WCDMA, and EDGE\(^6\) have driven wireless PA design toward methods to improve efficiency while maintaining linearity, both at full power (base stations with wide peak to average ratios) and at lower powers (for handsets, where battery life is a concern). Nothing has yet been published to explore applying Doherty’s concept to saturated amplifiers at backed-off power.

Many overview papers and book sections have been devoted to the analysis of the Doherty amplifier [15]-[19]. Raab’s important paper on the efficiency of Doherty systems [15] uncovered the efficiency relationship in the transition region, assuming two class B amplifiers. He also explored varying the power ratio between the two amplifiers, where the low power efficiency peak can be pushed more than 12 dB below full power. Called an extended Doherty amplifier, the value of \( \alpha = 0.25 \) is often used. The value of \( Z_{\text{INV}} \) for proper Doherty action then becomes a function of \( \alpha \),

\[
Z_{\text{INV}} = \frac{R_L}{\alpha}.
\]  

---

\(^6\) WCDMA: Wideband CDMA; EDGE: Enhanced Data-rate for GSM Evolution
Raab developed efficiency equations for the extended Doherty amplifier, starting with the standard class B efficiency equation (derived in Section 3.2),

$$\eta_{\text{CLASS B}} = \frac{\pi}{4} \frac{V_L}{V_{\text{LMAX}}} = \frac{\pi}{4} \frac{V_L}{4 V_{\text{CC}}}$$  \hspace{1cm} (10)$$

where $V_{\text{CC}}$ is the applied DC collector voltage, and $V_{\text{LMAX}}$ is the RF voltage magnitude across $R_L$ at full power, equal to $V_{\text{CC}}$ for an ideal transistor at class B saturation. Note that maximum efficiency is $\pi/4 = 78.5\%$ as shown in Fig. 1.2, and varies proportionately with $V_L$.

Below the transition point $Q_2$ is off, and $Q_1$ sees a constant impedance transformed by the impedance inverter. Thus the efficiency falls off below the transition point and depends on $\alpha$:

$$\eta = \eta_{Q1} = \frac{\pi}{4} \frac{V_1}{V_{\text{CC}}} = \frac{\pi}{4} \frac{V_L}{4 \alpha V_{\text{CC}}} \hspace{1cm} \text{for } 0 < V_L < \alpha V_{\text{CC}}$$  \hspace{1cm} (11)$$

where $V_L = \alpha V_{\text{CC}}$ defines the transition point. Since $Q_1$ is held at $V_1 = V_{\text{CC}}$ throughout the transition region, its efficiency is

$$\eta_{Q1} = \frac{\pi}{4} \hspace{1cm} \text{for } \alpha V_{\text{CC}} < V_L < V_{\text{CC}}.$$  \hspace{1cm} (12)$$

The peaking amplifier’s efficiency in the transition region, $\eta_{Q2}$, is identical to the class B curve. The composite efficiency will be some combination of $\eta_{Q1}$ and $\eta_{Q2}$, depending on which side is contributing more power at each point. Raab’s composite efficiency expression in the transition region is

$$\eta = \frac{\pi}{4} \frac{V_L / V_{\text{CC}}}{\alpha(1 - V_{\text{CC}} / V_L) + 1} \hspace{1cm} \text{for } \alpha V_{\text{CC}} < V_L < V_{\text{CC}}$$  \hspace{1cm} (13)$$
and is graphed versus power out in Fig. 2.5 for both $\alpha = 0.5$ and $\alpha = 0.25$. Note that the dip is more pronounced as $\alpha$ is made smaller, but the efficiency benefit at low power is greater.

![Ideal Doherty efficiency vs. power out.](image)

Raab further developed a multi-section Doherty amplifier as a way to reduce the dip in efficiency. The carrier amplifier is split into two (or more) amplifiers in a Doherty configuration, resulting in an efficiency curve with three (or more) peaks and less pronounced dips in efficiency. Multi-section Doherty amplifiers will not be considered in this thesis because of their increased size, complexity and cost.

More recently, there have been many papers that apply the Doherty concept to wireless power amplifiers, both base station and handset [3]-[14], [20]-[23], and at
microwave frequencies as high as K-band [24]-[26]. Many are hardware demonstration papers, most using quarter-wave transmission lines for tuning and impedance inverters, suitable for base station use. Due to nonlinear effects in the peaking amplifier [17], [5], considerable effort has been expended to improve phase and amplitude linearity in these Doherty amplifiers. Unfortunately, most of these solutions require a sacrifice in efficiency at lower power levels, and all hardware demonstrated for CDMA, WCDMA or multi-carrier applications has much lower efficiency at the transition point than at full power.

Fig. 2.6 shows a typical efficiency curve from this literature [8] that meets WCDMA base station linearity requirements. The PAE at 10 dB below full power is 20% lower, but this represents a great improvement in efficiency over the standard

![Fig. 2.6 Practical linear Doherty efficiency [8].](image-url)
class AB PA. Without the constraint of linearity, the low power peak in efficiency might approach that of the full power peak, as in Fig. 2.5.

This year Jonathan Lees reported on optimizing the Doherty amplifier at 1.8 GHz using single tone optimization, but still with an eye toward improving linear efficiency [13]. He advocates the use of adaptive gate bias, where the input envelope level is sensed or provided by the baseband processor, and the gate bias is adjusted to minimize quiescent current at each power level, while still meeting the linearity requirements. While he suggests the use of adaptive gate bias to control the conduction angle of the peaking amplifier, he also relies on a changing input drive level to achieve power control. He discusses some of the nonlinear challenges in Doherty design, such as phase mismatch between the two transistors and nonlinear phase delay as bias and drive levels are increased. He proposes the use of digital signal processing (DSP) to provide not only optimized adaptive bias but also RF input phase and magnitude correction for each amplifier to achieve high linear efficiency. Similar techniques could be used to optimize a saturated Doherty amplifier, possibly with much simpler and more appealing methods than a customized DSP solution.

2.3. Other Low Power Efficiency Enhancement Techniques

A promising technique currently under development for GSM handset PAs is to use high efficiency DC-DC converters combined with Vcc power control to keep the output device saturated at lower voltage swings. In a way, this is a parallel approach to the Doherty amplifier, where saturated voltage swings are maintained by lowering Vcc instead of dynamically raising the load impedance. A major challenge for this
approach is preventing RF spurious products, due to converter switching, from reaching the PA output.

There are several techniques used in CMDA handset PAs such as bias stepping, where at lower power levels the quiescent current is stepped back to raise efficiency while maintaining linearity, and switched periphery where part of the output transistor array is turned off at lower powers to further lower the quiescent current and raise the gain of the amplifier, improving PAE.

Higher class (D-F) switching PAs could be used to increase the efficiency at full and lower powers, but these are difficult to implement at wireless frequencies in a handset [17]. Reference [10] introduces a class-F Doherty amplifier combined with the feed forward technique for a base station application. It demonstrates higher maximum efficiency and an efficiency profile similar to that in Fig. 2.6. Higher class amplifiers are not used in this thesis because of the difficulty in filtering harmonics successfully enough to allow power combining (which requires pure sine waves at the point of combination). Complex harmonic terminations can introduce undesirable impedance transformations as the load impedance presented to these filters is changed through the Doherty action.

Many popular efficiency enhancement techniques, such as outphasing, feed forward, predistortion, Kahn and others are only applicable to linear modulation schemes [17], so are not within the scope of this thesis.
3. CIRCUIT ANALYSIS

This chapter will cover the “back of the napkin” analysis that should be done before the detailed circuit design begins. First, it is important to clearly define the primary goal for the final circuit: make cell phones last longer between battery charges. For a PA designer, the question becomes how to use less current so the milliamp-hours available from the battery go further. Armed with a power probability distribution curve as discussed in Section 1.1, it quickly becomes apparent that reducing current at lower power levels is just as important as at full power. So the problem then becomes an analysis of how to improve efficiency over the entire power range of the PA.

The primary way of controlling the efficiency of a PA is through the load that is presented to the final transistor. Section 3.2 will review the relationship between load impedance and class B efficiency. Next, load modification using Doherty action will be discussed, and then a novel solution to the Doherty circuit requirements will be presented in Sections 3.4 and 3.5. Finally, the new method will be analyzed over the GSM power PDF to see how much benefit in battery life can be obtained.

3.1. Problem Statement

It is very annoying when a cell phone stops working during a call, especially when it was recharged just a few days earlier. This annoyance turns into market pressure for cellular service providers, phone manufacturers, and ultimately for PA manufacturers. As mentioned in Chapter 1, GSM PA efficiency at full power has
reached a practical maximum in the industry. Therefore current reduction at lower power levels, if significant enough, is the main way to improve average current.

A recent study of GSM handset power usage statistics was performed by Miller Wireless Consulting [27]. Fig. 3.1 shows the distribution of transmit power levels for a handset in the 1900 MHz GSM band in an urban area. Other than the spike at full power, much time is spent at lower power, similar to the CDMA case discussed in Section 1.1. However, the current at lower power is also smaller than at full power, making it less significant to talk time. Fig. 3.2 shows the current for an ideal class B PA versus power output, scaled down to 50% collector efficiency at 33 dBm to represent the efficiency encountered in an normal GSM cell phone. Note that although the phone spends much of its time transmitting near a power out of 12 dBm

![Histogram of Mobile Transmit Power for Washington, DC Voice Calls](image)

Fig. 3.1 Urban PDF for GSM PCS1900 band [27].
in Fig. 3.1, the current consumed at that power is only 10% of that consumed at full power according to Fig. 3.2.

A figure of merit is needed to measure the overall improvement in talk time. A weighted average of current at each power level, using the PDF in Fig. 3.1 and currents from Fig. 3.2, would give a useful measure of talk time improvement. Indeed, if the PDF is a continuous function of power out [1],

\[
I_{C\_AVE} = \int_0^\infty [PDF(P_{OUT}) \cdot I_c(P_{OUT})] \, dP_{OUT} .
\]  

If the PDF is a discrete function as in Fig. 3.1, defined over the range of \( P_{OUT} = P_{MIN} \) to \( P_{MAX} \) and normalized such that \( \sum PDF = 1 \), then

\[
I_{C\_AVE} = \sum_{P_{MIN}}^{P_{MAX}} PDF(P_{OUT}) \cdot I_c(P_{OUT}) .
\]  

This summation is performed in Table 3.1. Since the power levels from the PDF in Fig. 3.1 are referenced at the antenna, 2 dB was added to the Pout column to
compensate for output network losses in the cell phone, yielding the power out at the 
PA. The PDF data was normalized to 1 (100%), then multiplied by the currents from 
Fig. 3.2. The summation totals 348 mA, which represents an average current for the 
PA over its operating power range. Note that the mean power of the PDF in Fig. 3.1 is 
16.2 dBm (18.2 dBm at the PA).

It is clear from looking at the partial sums in Table 3.1 that if low power 
efficiency can be improved, the average current can be reduced significantly. It will 
be shown in the following analysis that average current can be reduced by 30% with 
the proposed Doherty amplifier.

<table>
<thead>
<tr>
<th>Pout (dBm)</th>
<th>PDF (%)</th>
<th>Idc (A)</th>
<th>Partial sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>6.4%</td>
<td>0.032</td>
<td>0.002</td>
</tr>
<tr>
<td>4</td>
<td>2.4%</td>
<td>0.040</td>
<td>0.001</td>
</tr>
<tr>
<td>6</td>
<td>4.8%</td>
<td>0.051</td>
<td>0.002</td>
</tr>
<tr>
<td>8</td>
<td>5.1%</td>
<td>0.064</td>
<td>0.003</td>
</tr>
<tr>
<td>10</td>
<td>4.9%</td>
<td>0.081</td>
<td>0.004</td>
</tr>
<tr>
<td>12</td>
<td>8.2%</td>
<td>0.102</td>
<td>0.008</td>
</tr>
<tr>
<td>14</td>
<td>8.5%</td>
<td>0.128</td>
<td>0.011</td>
</tr>
<tr>
<td>16</td>
<td>6.3%</td>
<td>0.161</td>
<td>0.010</td>
</tr>
<tr>
<td>18</td>
<td>6.2%</td>
<td>0.203</td>
<td>0.013</td>
</tr>
<tr>
<td>20</td>
<td>7.5%</td>
<td>0.255</td>
<td>0.019</td>
</tr>
<tr>
<td>22</td>
<td>6.1%</td>
<td>0.321</td>
<td>0.019</td>
</tr>
<tr>
<td>24</td>
<td>5.9%</td>
<td>0.405</td>
<td>0.024</td>
</tr>
<tr>
<td>26</td>
<td>6.3%</td>
<td>0.509</td>
<td>0.032</td>
</tr>
<tr>
<td>28</td>
<td>3.4%</td>
<td>0.641</td>
<td>0.022</td>
</tr>
<tr>
<td>30</td>
<td>3.2%</td>
<td>0.807</td>
<td>0.026</td>
</tr>
<tr>
<td>32</td>
<td>14.9%</td>
<td>1.016</td>
<td>0.151</td>
</tr>
</tbody>
</table>

\[
I_{c\text{ave}} = 0.348
\]

Table 3.1 Summation from equation (15).
3.2. Optimizing Load Impedance for Efficiency

Before discussing how to optimize efficiency, a basic understanding of what makes a PA efficient is needed. A class B amplifier example will continue to be used because of its ease of analysis and its applicability to base power-controlled PAs.

Fig. 3.3 shows the class B collector waveforms for voltage and current at saturation [16]. Recall from Section 1.3 that a class B amplifier requires all harmonics to be shorted at the output, therefore shunting all harmonic currents to ground and leaving only the fundamental frequency component of current to pass through the load. This forces the voltage waveform to be a pure sine wave with magnitude \( V_L \), as shown.

![Class B waveforms](image)

**Fig. 3.3** Class B waveforms.
Collector efficiency is defined as the ratio of RF power delivered to DC power consumed, or

\[
\eta = \frac{P_{\text{OUT}}}{P_{\text{DC}}} = \frac{\left(\frac{I_{f1}}{\sqrt{2}}\right)^2 R_L}{I_{DC} V_{CC}} = \frac{I_{f1}^2 R_L}{2I_{DC} V_{CC}}
\]

where \(I_{f1}\) is the magnitude of the fundamental component of the collector current, \(R_L\) is the load resistance and \(I_{DC}\) is the DC component of the collector current waveform.

Since the current waveform is a half sine wave, its DC and fundamental values from the Fourier Series are

\[
I_{DC} = \frac{I_{\text{PEAK}}}{\pi}
\]

\[
I_{f1} = \frac{I_{\text{PEAK}}}{2}
\]

Substituting these into (16) gives

\[
\eta = \frac{I_{f1} I_{\text{PEAK}} R_L}{4 \frac{\pi}{I_{\text{PEAK}} V_{CC}}} = \frac{\pi I_{f1} R_L}{4 V_{CC}}
\]

\[
\eta = \frac{\pi V_L}{4 V_{CC}}
\]

which confirms (10) in Section 2.2. At saturation, \(V_L = V_{CC}\) and \(\eta = \pi/4\) or 78.5%, as expected. The voltage waveform in Fig. 3.3 shows this full voltage swing, where \(V_C\) ranges from 0 to 2\(V_{CC}\).

To get this saturation condition, the right load impedance is needed. If \(I_{f1}\) is assumed constant, which would be the case for fixed input drive, then \(V_L = I_{f1} R_L\) dictates that \(R_L\) alone sets \(V_L\). Maximum power output for a given \(I_{f1}\) also occurs at
the saturation point, so the key to optimum load impedance is derived from the simple power equation

\[ P_{\text{OUT,W}}(\text{sat}) = \frac{V_L^2}{R_L} = \frac{V_{L_{\text{sat}}}^2}{2R_L} = \frac{V_{\text{cc}}^2}{2R_{L_{\text{OPT}}}} \]  

(21)

where \(P_{\text{OUT,W}}\) is the output power expressed in watts, \(V_{\text{cc}}\) represents the magnitude of \(V_L\) at saturation, and \(R_{L_{\text{OPT}}}\) is the minimum load resistance needed to reach saturation at a given \(P_{\text{OUT,W}}\). Solving (21) for \(R_{L_{\text{OPT}}}\) leads to

\[ R_{L_{\text{OPT}}} = \frac{V_{\text{cc}}^2}{2P_{\text{OUT,W}}} \]  

(22)

Now if \(I_{f1}\) is allowed to vary (from a changing input drive due to power control), there will be an optimum load resistance that will maximize efficiency (and power) at each \(I_{f1}\). Fig. 3.4 shows the relationship between \(R_{L_{\text{OPT}}}\) and \(P_{\text{OUT}}\) using (22) and

Fig. 3.4  Optimum class B load impedance vs. power out, \(V_{\text{cc}} = 3.5\) V.
assuming that $V_{CC} = 3.5$ V. Note that a perfect efficiency of 78.5% could theoretically be maintained over the entire power range if the load resistance was dynamically varied from 3 $\Omega$ to 6125 $\Omega$.

For a normal class B amplifier without load modulation, the maximum power required by the system will determine the load resistance according to (22). As a lower power is demanded by the system and the input drive is reduced, $I_f$ will also be reduced, resulting in a lower voltage swing as shown by the dashed line in Fig. 3.5. Efficiency at this lower power also reduces according to (20).

![Efficiency diagram](image.png)

Fig. 3.5 Reduced output power waveforms and power dissipation.

Efficiency is directly related to the DC power dissipated in the transistor. One way to visualize this power loss is to imagine integrating the instantaneous voltage-current product ($I_C \cdot V_C$) across the transistor over an RF cycle. Since the current is
zero for half the conduction cycle, the only opportunity for instantaneous power dissipation in the transistor is during the lower half cycle of the voltage swing. Fig. 3.5 shows this region of power dissipation as a red hashed area, where the product of $I_C$ and $V_C$ is non-zero. The total loss during an RF cycle can be visualized as integrating the “area under the curve” of the I-V product in this region. Obviously as the voltage swing is reduced, the minimum voltage in the voltage waveform rises above zero, and the “area under the curve” of the I-V product goes up markedly.

Raising the load impedance at lower drive levels has the effect of increasing the voltage swing and raising the efficiency, as in (19). In order to maximize efficiency as the drive level is changed, a dynamically changing load impedance is needed to maintain saturation in the face of a changing $I_{f1}$.

3.3. Doherty Load Modulation

This dynamic load requirement is achieved in the Doherty amplifier, at least for Q1 in Fig. 3.7. When the rather rigorous input drive relationships for the two amplifiers are met, Q1 is kept at the edge of saturation throughout the transition region, as discussed in Section 2.1 (Doherty’s requirements from Fig. 2.3 are repeated below in Fig. 3.6 for convenience). Under these conditions, the load resistance presented to Q1 is simply the RF voltage across device 1 from Fig. 3.6a, divided by the device 1 output current from Fig. 3.6b. The voltage is assumed to be a straight line equal to $V_{CC}$ throughout the transition region, so

$$R_i = \frac{V_i}{I_i} = \frac{V_{CC}}{I_1} \quad for \quad \alpha V_{CC} < V_L < V_{CC}, \quad (23)$$
where $V_L$ is the output voltage, $\alpha V_{CC}$ is the output voltage at the transition point, and $V_L = V_{CC}$ represents full power out. From (4) and (9) (repeated below),

$$I_1 = \frac{V_L}{Z_{INV}}$$  \hspace{1cm} (4)
\[ Z_{\text{INV}} = \frac{R_L}{\alpha} \]  \hspace{1cm} (9)

\[ I_1 = \frac{V_L}{Z_{\text{INV}}} = \frac{\alpha}{R_L} V_L \]  \hspace{1cm} (24)

\[ R_1 = \frac{V_{cc}}{I_1} = V_{cc}\left(\frac{\alpha}{R_L} V_L\right) = \frac{V_{cc} R_L}{\alpha} \frac{1}{V_L} \quad \text{for } \alpha V_{cc} < V_L < V_{cc} \]  \hspace{1cm} (25)

which is an inverse function of \( V_L \). \( P_1 \), the power delivered from Q1 is

\[ P_1 = \frac{V_L}{\sqrt{2}} I_1 = \frac{V_L I_1}{2} = \frac{V_{cc}}{2} \frac{\alpha}{R_L} V_L \quad \text{for } \alpha V_{cc} < V_L < V_{cc} . \]  \hspace{1cm} (26)

Using (25) and (26), \( R_1 \) can be plotted in the transition region and compared to the results from Fig. 3.4, as shown in Fig. 3.8. Note that the ideal class B Doherty

![Graph](image)

Fig. 3.8  Load presented to Q1 in Doherty amplifier, \( V_{cc} = 3.5 \) V, \( \alpha = 0.25 \).
amplifier provides the optimum load for maximum class B efficiency throughout the transition region.

Below the transition point Q2 is off, so \( R_{IT} = R_L \). This impedance is transformed through \( Z_{INV} \) to a constant value by (3) and (9) (repeated below):

\[
R_I R_{IT} = Z_{INV}^2
\]

(3)

\[
Z_{INV} = \frac{R_L}{\alpha}
\]

(9)

leading to

\[
R_I = \frac{Z_{INV}^2}{R_{IT}} = \left( \frac{R_L}{\alpha} \right)^2 \frac{1}{R_L} = \frac{R_L}{\alpha^2} \quad \text{for} \quad 0 < V_L < aV_{CC}.
\]

(26)

This value for \( R_I \) below the transition point has been added to Fig. 3.8 and shows the impedance stabilized at a relatively high level for all powers below the transition point.

The efficiency of Q2 has no such benefit from a modulated load impedance, but it turns out its efficiency is not worsened, either. As shown in (20) and visualized in Fig. 3.5, the efficiency of Q2 is determined only by the output voltage amplitude across Q2. Some of the voltage swing across the load is due to the current from Q1, so the efficiency of Q2 is a function of the total power output of both amplifiers, not just of Q2:

\[
\eta_{Q2} = \frac{\pi V_I}{4 V_{CC}} = \eta_{CLASS B} (total \ power \ out).
\]

(27)

The total amplifier efficiency is a composite of the two amplifiers’ efficiencies, as shown in Fig. 2.5 for \( \alpha = 0.25 \). At full power and for this \( \alpha \), Q2 produces three
times more power than Q1, so the efficiency there is dominated by Q2’s efficiency. As power out is reduced, Q2’s power drops more quickly than Q1’s, until it is finally zero at the transition point, 12 dB below full power. Therefore Q2’s efficiency continues to dominate the composite efficiency until its power out is less than Q1’s, which explains why the dip is deeper for small values of $\alpha$, and efficiency only recovers near the lower peak.

The depth of the dip is an important factor to talk time. From the earlier discussion of weighed average current (as in (15) and Table 3.1), the strongest contributors to average current occur within 8 dB of full power (see the “partial sum” column of Table 3.1). Therefore to improve the composite efficiency, Q2’s efficiency must be improved to be better than the class B curve, at least near full power.

To achieve this, it is desirable to somehow trade in some of Q2’s linearity for more efficiency. One way would be to bias Q2 as a class C amplifier. Referring to Fig. 3.5, reducing the conduction angle would reduce the “area under the curve” product of current and voltage across the device, and reduce the dissipation in the transistor. In fact, if a method could be devised to reduce the conduction angle of Q2 (and Q1, if possible) continuously from 180° at full power to zero at the transition point, higher average efficiency should result. It will be shown in the next section that power control by adjusting the base voltage gives this result.

3.4. The Base Control Amplifier

Base control has been used as a method of controlling the output power of GSM handset PAs for several years. A power control voltage, $V_{PC}$, is provided to the PA
and controls the bias point of the bases of all stages (see Fig. 1.4 in Section 1.2 for a block diagram of the base controlled power control loop). The RF input power to the PA is fixed. As $V_{PC}$ is lowered from its maximum, the magnitude and conduction angle of the drive signal at the bases is reduced, resulting in lower output power and improved efficiency due to the smaller conduction angle. The operating point of each transistor is always at its minimum current for a given power level, so at low power levels there is no excess in quiescent current which would otherwise lead to lower efficiency in a class AB amplifier.

Using base control in a Doherty amplifier makes sense not only because of the improved efficiency, but also to provide a convenient way to control the drive relationship required for proper Doherty action. This will be discussed in more detail in Section 3.5.

A two-stage base controlled PA is shown in Fig. 3.9a, and its equivalent circuit is shown in Fig. 3.9b, assuming the following simplifications:

- Only the output stage (Q2) is analyzed, and the drive from the input stage (Q_{DRIVER}) is assumed constant, even though it is also base controlled. This assumption works well in practice because the $V_{PC}$ applied to Q2 primarily controls the higher power region, whereas the drive level to Q2 (or $V_{PC}$ to Q_{DRIVER}) dominates the control at lower powers. A comparison to data in Section 4.5 will further support this assumption.

- Q_{DRIVER} is further assumed to act like a constant RF current source that drives the base of Q2. An interstage matching network is usually used to transform
the $Q_{\text{DRIVER}}$ output impedance to the $Q2$ input impedance, providing a source impedance that is matched to the base. To simplify the analysis, a constant current approximation (high source impedance) appears to be more accurate than a voltage source approximation. The classic analysis of ideal transconductance devices and voltage drive at the base seems to breakdown.

Fig. 3.9 Proposed base control circuit and its equivalent circuit for analysis.
beyond repair when applied to bipolar transistors operating several dB into compression. The transistor in this analysis will instead be treated as a linear current amplifier.

- The harmonic currents from the first stage and second stage are assumed to be terminated perfectly, as is typically assumed for class B and class C operation. The parallel resonator in Fig. 3.9 is tuned to the fundamental RF frequency and is assumed to have a high Q.

- The Q2 base-emitter junction is assumed to be an ideal diode, with a built in voltage of \( V_{BI} = 1.2 \) V, to approximate the GaAs HBT technology that will be used later in the design. The ideal diode is defined here as a switch that is an open circuit below 1.2 V, and is a short circuit above 1.2 V. Furthermore, the HBT is assumed to be a linear current amplifier, where the collector current is equal to the current through the ideal diode multiplied by the RF beta of the transistor.

The circuit in Fig. 3.9b shows current waveforms at full power assuming class B operation. A half sine from the class B driver stage is turned into a sinusoidal current by the harmonic trap. Enough DC current is contributed by the \( V_{PC} \) supply to allow the diode to conduct for half of the RF cycle (class B drive). In fact, the RF current that is conducted through the diode is delivered from the blocking capacitor, and when the diode is off, the \( V_{PC} \) supply absorbs the negative half cycle and “recharges” the blocking capacitor. The result is a half sine wave of base current that is multiplied by beta and reproduced at the collector.
Fig. 3.10 on the next page shows what happens as the $V_{PC}$ voltage is changed and the amplifier is reduced from class B conduction at full power, to class C at lower power, and finally to cutoff. Both amplitude and conduction angle of the base current are reduced as $V_{PC}$ is reduced, a better situation for efficiency than in the case of backed off class B, where the conduction angle remains 180°. The lower conduction angle of the output current waveform will provide a modest improvement in efficiency as the “area under the curve” in Fig. 3.5 is reduced somewhat. Note also that the top of the base voltage waveform is flattened due to diode conduction above 1.2 V, an effect often seen in measured bipolar PA waveforms.
Fig. 3.10  Base current and voltage waveforms as $V_{PC}$ is reduced.
Fig. 3.11 shows a detailed view of the collector current waveforms as $V_{PC}$ is reduced. A cosine waveform is used in Fig. 3.11 to easily see that conduction angle, $\gamma$, is reducing as $V_{PC}$ is reduced. The three cases represented in Fig. 3.10 are shown: the case in Fig 3.10a is represented by the waveform with $I_{MAX1}$ as its maximum value, Fig 3.10b by the curve with maximum value of $I_{MAX2}$, and 3.10c by the curve with $I_{MAX3} = 0$. $I_{PK}$ is the magnitude of the sine wave and is constant. $I_{Q}$ is the “zero-crossing” level for the sine wave, and becomes more negative as $V_{PC}$ is reduced. $I_{MAX}$ represents the maximum current with respect to zero for a given $I_{PK}$ and $I_{Q}$, so that

$$I_{PK} = I_{MAX} - I_{Q}.$$ (28)

Q2’s efficiency can be derived by modifying Cripps’ analysis for reduced conduction angle PAs [16, p48]. From Fourier transform analysis, the fundamental and DC components of the collector current waveforms, using the notation from Fig. 3.11, are

Fig. 3.11  Base control collector current waveforms.
\[ I_{DC} = I_{\text{MAX}} \frac{\sin(\gamma/2) - (\gamma/2) \cos(\gamma/2)}{1 - \cos(\gamma/2)} \]  

(29)

\[ I_1 = I_{\text{MAX}} \frac{\gamma - \sin \gamma}{2\pi} \frac{1}{1 - \cos(\gamma/2)} \]  

(30)

where \( \gamma \) is the conduction angle, and \( \gamma = \pi \) for class B operation. The waveform in Fig. 3.11 is a cosine function, so the relationship between the angle \( \gamma \) and \( I_Q \) is determined from by

\[ I_{PK} \cos(\gamma/2) = -I_Q \]  

(31)

since \( I_Q \) is defined to be negative for \( \gamma > 0 \). Substituting for \( I_Q \) from (28),

\[ I_{PK} \cos(\gamma/2) = I_{PK} - I_{\text{MAX}} \]  

(32)

\[ I_{\text{MAX}} = I_{PK} [1 - \cos(\gamma/2)] . \]  

(33)

At full power out, \( I_{\text{MAX}} = I_{PK} \), and the waveform is a half sine wave and operating in class B mode. The output load resistance, \( R_L \), is assumed to be chosen such that the voltage swing is saturated at full power. So the fundamental Fourier component of the waveform is

\[ I_1(\text{full power}) = I_{\text{MAX}}(\text{full power}) = \frac{I_{PK}}{2} \]  

(34)

\[ I_{PK} = 2I_1(\text{full power}) = 2 \left( \frac{V_L(\text{full power})}{R_L} \right) = 2 \left( \frac{V_{CC}}{R_L} \right) . \]  

(35)
Substituting (35) into (33), and this result into (29) and (30) gives

\[ I_{MAX} = \frac{2V_{CC}}{R_L} [1 - \cos(\gamma/2)] \]  
(36)

\[ I_{DC} = \frac{2V_{CC}}{\pi R_L} [\sin(\gamma/2) - (\gamma/2)\cos(\gamma/2)] \]  
(37)

\[ I_1 = \frac{V_{CC}}{\pi R_L} (\gamma - \sin \gamma) . \]  
(38)

Deriving \( V_L \) in terms of \( \gamma \) using (38),

\[ V_L = I_1 R_L = R_L \left[ \frac{V_{CC}}{\pi R_L} (\gamma - \sin \gamma) \right] \]  
(39)

\[ V_L = \frac{V_{CC}}{\pi} (\gamma - \sin \gamma) . \]  
(40)

Finally, collector efficiency for the base controlled transistor, \( \eta_{BC} \), is derived using the results for \( V_L, I_1, \) and \( I_{DC} \) above:

\[ \eta_{BC} = \frac{P_{OUT}}{P_{DC}} = \frac{V_L I_1/2}{V_{CC} I_{DC}} \]  
(41)

\[ \eta_{BC} = \frac{1}{2V_{CC}} \left[ \frac{V_{CC}}{\pi} (\gamma - \sin \gamma) \right] \cdot \frac{\frac{V_{CC}}{\pi R_L} (\gamma - \sin \gamma)}{\frac{2V_{CC}}{\pi R_L} [\sin(\gamma/2) - (\gamma/2)\cos(\gamma/2)]} \]  
(42)

\[ \eta_{BC} = \frac{1}{4\pi} \cdot \frac{(\gamma - \sin \gamma)^2}{\sin(\gamma/2) - (\gamma/2)\cos(\gamma/2)} \]  
(43a)

\[ V_L = \frac{V_{CC}}{\pi} (\gamma - \sin \gamma) . \]  
(43b)
Equations (43a) and (43b) can be evaluated together and the efficiency plotted versus power out. Fig. 3.12 shows this plot compared to the ideal class B curve derived earlier. Note that efficiency is somewhat improved due to the smaller conduction angle, around 2% at 1 dB back off from full power, and 4% at 3 dB back off.

![Efficiency vs. Power Out](image)

**Fig. 3.12** Base control efficiency vs. power out.

### 3.5. Applying Base Control to the Doherty Amplifier

In this section, base control will be combined with the Doherty concepts developed earlier. If used for constant envelope signals, several advantages over more traditional Doherty implementations can be realized with this approach.
Getting the right balance of output currents through the transition region (as in Fig. 3.6) has been a weakness of the Doherty amplifier since its inception. Linear Doherty implementations depend on carefully balancing the RF drive levels between the two amplifiers by one of three methods:

- biasing Q2 (in Fig. 3.7) for class C operation so its large signal gain increases more rapidly than the class B or AB Q1;
- dynamically adjusting the bias point of Q2 (adaptive bias) by detecting the input power level or receiving this information from the baseband controller; or
- separately adjusting the RF input levels at the inputs of Q1 and Q2, either by providing separate RF paths from the transceiver IC or with attenuators or variable gain amplifiers (VGA) at the PA.

Implementing a Doherty amplifier using base control to get the correct balance of output currents has several advantages over these other methods:

- It is an effective and flexible way to optimize the drive levels for Doherty action. DC voltages (or currents) can be provided to each amplifier, or even each stage if desired, to deliver the appropriate drive requirements in Fig. 3.6 at each power level. An analog circuit on the amplifier die could be used, or for more precise control, a custom silicon controller could be added to the PA module. The small current required from a silicon controller would mean the die could be small and manufactured at low cost in the volumes typically seen for cellular telephones.
• It provides for more backed off efficiency, in both Q2 and Q1, than the
typical class B or class AB amplifiers used in linear Doherty amplifiers. As
discussed earlier, the conduction angle reduces from class B at full power to
backed off class C, for a significant improvement in efficiency over a class B
amplifier. A base controlled Doherty amplifier effectively trades off the
linearity of the RF transfer function for more for efficiency for constant
envelope signals.

• Due to the simplicity of the circuit, it is cost effective when used in a coupler-
   based power control loop.

A base control Doherty amplifier circuit is the same as the traditional Doherty
structure shown in Fig. 3.7 except that the RF input power is kept constant and the
base bias point is adjusted independently for each amplifier, as discussed in the
previous section. Using the idealized circuit from Fig. 3.9 and Fig. 3.10, separate $V_{PC}$
voltages for each amplifier can be used to directly produce the correctly balanced RF
collector currents. Although the non-linear junction voltage in a real transistor
complicates the picture somewhat, it will be demonstrated in Section 4.2 that proper
Doherty action can still be attained without altering this simple approach.
Using the base control Doherty concept efficiency is calculated again using (43a) and (43b) instead of the class B efficiency equation. Fig. 3.13 shows this new curve compared to the class B Doherty amplifier from Fig. 2.5.

Note that the efficiency has increased significantly, especially at the transition point. In the transition region, Q1 is operating in true class C mode: as the conduction angle is reduced, the Doherty action is keeping Q1 in saturation, thus at the maximum class C efficiency for the resulting conduction angle.

Of course, the efficiency of the final circuit will be reduced when the current from the driver stages, losses in the matching network, and non-ideal transistor effects are taken into account. Nevertheless, the improvement in efficiency near the transition
point and throughout the transition region should have a significant effect on the average PA current, resulting in improved talk time.

3.6. Talk Time Analysis

When the results from the last section are applied to the PDF curve in Fig. 3.1, a substantial improvement in average PA current results. Table 3.2 repeats the calculations performed in Table 3.1 (Section 3.1) for the class B Doherty and base control Doherty PAs, again with all calculations scaled to 50% efficiency at full power. The weighted average current for the Class B Doherty PA decreases by 29% over the class B amplifier, and the base control Doherty improves by 34%.

<table>
<thead>
<tr>
<th>Pout (dBm)</th>
<th>Class B Amplifier</th>
<th>Class B Doherty</th>
<th>Base Ctrl Doherty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pout (%)</td>
<td>Idc (A)</td>
<td>Partial sum</td>
</tr>
<tr>
<td>2</td>
<td>6.4%</td>
<td>0.032</td>
<td>0.002</td>
</tr>
<tr>
<td>4</td>
<td>2.4%</td>
<td>0.040</td>
<td>0.001</td>
</tr>
<tr>
<td>6</td>
<td>4.8%</td>
<td>0.051</td>
<td>0.002</td>
</tr>
<tr>
<td>8</td>
<td>5.1%</td>
<td>0.064</td>
<td>0.003</td>
</tr>
<tr>
<td>10</td>
<td>4.9%</td>
<td>0.081</td>
<td>0.004</td>
</tr>
<tr>
<td>12</td>
<td>8.2%</td>
<td>0.102</td>
<td>0.008</td>
</tr>
<tr>
<td>14</td>
<td>8.5%</td>
<td>0.128</td>
<td>0.011</td>
</tr>
<tr>
<td>16</td>
<td>6.3%</td>
<td>0.161</td>
<td>0.010</td>
</tr>
<tr>
<td>18</td>
<td>6.2%</td>
<td>0.203</td>
<td>0.013</td>
</tr>
<tr>
<td>20</td>
<td>7.5%</td>
<td>0.255</td>
<td>0.019</td>
</tr>
<tr>
<td>22</td>
<td>6.1%</td>
<td>0.321</td>
<td>0.019</td>
</tr>
<tr>
<td>24</td>
<td>5.9%</td>
<td>0.405</td>
<td>0.024</td>
</tr>
<tr>
<td>26</td>
<td>6.3%</td>
<td>0.509</td>
<td>0.032</td>
</tr>
<tr>
<td>28</td>
<td>3.4%</td>
<td>0.641</td>
<td>0.022</td>
</tr>
<tr>
<td>30</td>
<td>3.2%</td>
<td>0.807</td>
<td>0.026</td>
</tr>
<tr>
<td>32</td>
<td>14.9%</td>
<td>1.016</td>
<td>0.151</td>
</tr>
</tbody>
</table>

Ic_ave = 0.348 | Ic_ave = 0.248 | Ic_ave = 0.230

Table 3.2 Improvement in average PA current.
To calculate the improvement in talk time, the battery capacity in mA-hours and the total phone current during a call must be determined. The average PA current over its operating power range is known, so a value for the average current consumed by the rest of the cell phone is needed. This was recently measured by Chris Stephens at TriQuint Semiconductor [28]. The non-PA current, averaged over the entire transmit frame was 97.4 mA, and the battery capacity was rated at 1000 mA-hours. The talk time can then be calculated as

\[ t_{\text{TALK}} = 60 \frac{B_{\text{CAP}}}{I_{\text{TOTAL}}} = \frac{60B_{\text{CAP}}}{rI_{\text{PA}} + I_{\text{NON-PA}}} \]

where \( t_{\text{TALK}} \) is the talk time in minutes, \( B_{\text{CAP}} \) is the battery capacity in mA-hours, \( r \) is the duty cycle ratio, and \( I_{\text{PA}} \) and \( I_{\text{NON-PA}} \) are in mA. During a voice call the PA is operated at 12.5% duty cycle, and during a GSM data transmission the duty cycle can be as high as 50%. Table 3.3 shows the results of the talk time calculations for 12.5% and 50% duty cycle, compared to the class B amplifier.

<table>
<thead>
<tr>
<th>Scaled to ( \eta = 50% )</th>
<th>Improvement in PA current</th>
<th>Improvement in talk time</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% less current</td>
<td>12.5% duty cycle</td>
</tr>
<tr>
<td>Class B Amplifier</td>
<td>348</td>
<td>--</td>
</tr>
<tr>
<td>Class B Doherty</td>
<td>248</td>
<td>29%</td>
</tr>
<tr>
<td>Base Ctrl Doherty</td>
<td>230</td>
<td>34%</td>
</tr>
</tbody>
</table>

Table 3.3 Summary of average PA current and talk time improvement.
Talk time improved 42 minutes or 10% for the class B Doherty and 50 minutes or 12% for the base control Doherty. This is a smaller talk time improvement than was hoped for, but is explained by comparing the average PA current to the non-PA current, with duty cycle taken into account. At 12.5% duty cycle, the class B PA’s average current of 348 mA becomes $348/8 = 44$ mA compared to the 97 mA for the rest of the phone. At 50% duty cycle, the PA current increases to $348/2 = 174$ mA, which explains why the talk time improvement is more dramatic for the 50% duty cycle case. It is deceptive to think of the PA as drawing 1.1 A at full power, when for talk time purposes this current must be reduced because of the effects of both duty cycle and the power PDF. Still, 50 minutes of additional talk time is a significant improvement. Also, the benefit to talk time may be greater in the future as current consumption in the rest of the phone is reduced and 50% duty cycle data calls become more prevalent.
4. PROTOTYPE DESIGN

In this chapter, a 30 MHz base control Doherty amplifier is designed and then simulated. First the amplifier topology is chosen and then single-stage stability is evaluated for the HBT device to be used. Then each section of the design is reviewed in detail. The final prototype design is then simulated and optimized for phase matching at the power combining point (see Fig. 4.5 at the end of Section 4.4 for the complete circuit).

The emphasis was placed on a simple design that demonstrates the basic benefits of this circuit concept. The design frequency of 30 MHz was chosen to allow simpler circuit implementation, and to allow the circuit to be built more easily with discrete components. This approach dramatically reduces the time required for design changes and experiments. What might have taken a month-long die spin would take an afternoon. Working at low frequency also makes it possible to examine voltage waveforms in the hardware circuit and compare these to simulation, aiding the troubleshooting of a problem. The main drawback to working at 30 MHz is that distributed element tuning is ruled out by the ten meter wavelength.

4.1. Topology and Device Technology

Two amplifier stages are needed to provide enough gain for a competitive GSM PA (~30 dB). Given this, it is possible to (a) design both halves of the Doherty amplifier as two-stage amplifiers (as in Fig. 1.3), with impedance inverters at their inputs and outputs, or (b) put the first impedance inverter after the first stage, so the
Doherty amplifier is composed of only the two second-stage transistors, with the second impedance inverter at their outputs.

Option (a) was chosen to make the input impedance inverter more stable, anchored by the 50 Ω source impedance. If the impedance inverter were implemented in the interstage as in option (b), the impedances on both sides of the inverter would change dynamically as power is swept, making for a highly constrained interstage matching problem. Approach (a) also allows the first stage transistor of PA #1 to be smaller in area, achieving higher power added efficiency for PA #1. Furthermore, implementing the impedance inverter at the 50 Ω impedance point allows lower Q components to be used, and opens up the possibility that it can be combined with the input matching networks, since the impedance inversion function is not needed (or even desirable), but only the 90 degree phase shift is important. Finally, this option allows the input impedance inverter to be easily placed off-chip, on the module substrate, where a printed inductor could be used to save cost.

The power split between the two halves of the Doherty amplifier was chosen to be $\alpha = 0.25$, which implies PA #1 will provide a fourth of the total power at full power out and the low power efficiency peak will be 12 dB below the full power peak. Experiments show that this is nearly the best choice for $\alpha$ to optimize the talk time improvement, given the class B Doherty efficiency profile, power PDF and 12.5% duty cycle for voice calls ($\alpha = 0.32$ was found to be optimum).

The device technology chosen was GaAs InGaP HBTs from TriQuint Semiconductor (TriQuint foundry process TQHBT). GaAs HBTs are currently the
dominant device technology for GSM handset PAs, and this choice ties in well with the author’s current work (the simulator model integrity is known to be good from past experience). A lower frequency discrete device was not used because good, correlated simulation models were not readily available, and designing with HBTs at 30 MHz offers the advantage of transitioning to 1900 MHz in steps, using the same devices and even transistor arrays if desired. It turned out that after stabilizing the HBTs at 30 MHz, their gain was close to that at 1900 MHz.

Before starting the design process, the size (number of standard transistor cells, see Fig 4.1) of each of the four stages was determined by the maximum power it would need to deliver. Then a single transistor was set up for S-parameter

![Basic HBT cell structure. Each transistor, capacitor and resistor forms a cell.](image-url)

Fig. 4.1 Basic HBT cell structure. Each transistor, capacitor and resistor forms a cell.
simulations, and the stability factor was calculated. Due to the extremely high gain at 30 MHz for these devices, they were very unstable without feedback of some sort. After experimenting with different methods, a combination of base resistance and base-collector feedback was adjusted until the stability factor was greater than 1 for all frequencies. The basic cell structure commonly used for HBT PAs, shown in Fig. 4.1, uses a resistor to apply bias to the base of each cell to “ballast” the devices, preventing current hogging and thermal runaway. A capacitor is used to couple the RF input to the base of each cell. It was found that the resistor value commonly used for ballasting was nearly enough to stabilize the HBT and reduce the gain to a reasonable level. With the addition of base to collector feedback, the transistor circuits were unconditionally stable and similar to the stable gain experienced at 1900 MHz. So for all transistors in this design, both RF and bias are fed to the BIAS port in Fig. 4.1, and the blocking capacitors are not used.

The basic circuit topology that was chosen is shown in the simplified schematic in Fig. 4.2. To facilitate the following discussion, the PA #1 or carrier amplifier will be called the low power amplifier, and PA #2 or the peaking amplifier will be referred to as the high power amplifier.
4.2. Bias Circuits

In order to get proper Doherty action, the fundamental currents in Q2\textsubscript{LP} and Q2\textsubscript{HP} (Fig. 4.2) must obey the linear relationships put forward by Doherty (see Fig. 2.3b, where $\alpha = 0.5$). Fig. 4.3 shows these ideal relationships for $\alpha = 0.25$ as the output voltage, $V_L$, is increased. The $I_{LP}$ line represents the amount of fundamental current from Q2\textsubscript{LP}, and its starting point is at $V_L = 0$ (zero power out). $I_{HP}$ represents the Q2\textsubscript{HP} fundamental current, and its starting point is at the transition point, or

\begin{equation}
V_L (\text{transition pt}) = \alpha V_{CC} = 0.25 V_{CC}
\end{equation}

\begin{equation}
I_{LP} (\text{transition pt}) = \alpha^2 I_L (P_{MAX}) = 0.0625 I_L (P_{MAX})
\end{equation}

Fig. 4.2 Prototype circuit topology.
where $I_L$ is the total fundamental current to the load, as shown in Fig. 4.2. The end
points for each of these lines are at full power ($P_{MAX}$), and are calculated from

$$I_{LP}(P_{MAX}) = \alpha I_L(P_{MAX}) = 0.25I_L(P_{MAX}) \quad (46)$$

$$I_{HP}(P_{MAX}) = (1-\alpha)I_L(P_{MAX}) = 0.75I_L(P_{MAX}) \quad (47)$$

Fig. 4.3 Doherty fundamental current relationships for $\alpha = 0.25$.

To approximate the relationships in Fig. 4.3, a control voltage is applied through
a resistor, as demonstrated in Section 3.4 (Fig. 3.9a), to create an increasing DC base
current as the control voltage is increased. Unlike the earlier analysis (Section 3.4
used an ideal base-emitter diode with zero on resistance), the resulting collector
current will not be entirely linear with respect to the applied control voltage because
the diode “on” resistance varies with base current. But it is a reasonable simplification that is easy to implement in hardware and approximates the curves in Fig. 4.3 with respect to $V_L$ very well. Further work is needed to develop an analog bias control IC that will more completely satisfy the relationship with respect to $V_{PC}$ in Fig. 4.3.

From simulation experiments, at $V_{PC} = 0.9$ V and with an RF drive of 0.3 V to the base, $Q_{2LP}$ is at the threshold of cutoff as in Fig. 3.10c. This defines the starting $V_{PC}$ for the $I_{LP}$ curve, and $V_{PC} = 2.2$ V was chosen for its end point. The $V_{PC}$ starting point for the $I_{HP}$ curve can be determined by adding the $V_{PC}$ axis to Fig. 4.3, and scaling $V_{PC}$ for $V_L = \alpha V_{CC}$, which results in $V_{PC} = 1.225$ V for $Q_{2HP}$ cutoff, as shown. This implies that if the $I_{HP}$ curve could start at a $V_{PC}$ of 0.325 V later than the $I_{LP}$ curve, then good Doherty action might be approximated. An “offset” voltage of $V_{OFFSET} = 0.325$ was implemented in the circuit with an additional power supply (this offset would be implemented in an analog bias control IC in the final application).

Next the proper values of resistors were chosen to give the final $I_{LP}$ and $I_{HP}$ currents necessary at $V_{PC} = 2.2$ V, completing the bias circuit. Fig. 4.4 shows simulation results after the design was completed, and can be compared to Fig. 4.3. The upper graph of fundamental currents versus $V_L$ matches the ideal curves in Fig. 4.3 very well. In the lower graph, $Q_{2LP}$ was found to turn on at a lower $V_{PC}$ than expected ($V_{PC} = 0.6$ V instead of 0.9 V). This appears to be the result of uneven drive levels between the two sides. Since $Q_{1LP}$ has a higher input impedance than $Q_{1HP}$, the higher voltage swing at the base of $Q_{1LP}$ causes it to turn on at a lower $V_{PC}$ than predicted. As a result, $Q_{2LP}$ saturates at $V_{PC} = 0.8$ V, and $I_{LP}$ “stalls” at that level until
Q2\textsubscript{HP} turns on. When V\textsubscript{PC} rises to 1.2 V, Q2\textsubscript{HP} turns on and Q2\textsubscript{LP} continues its upward current trend as its load is modulated to a lower value.

Although the “stall” in the power ramp is undesirable in the final application, the circuit adequately shows the benefits of the base control Doherty amplifier. Future work would remove this “stalled” region by either modifying the input matching networks and/or decreasing V\textsubscript{OFFSET}.

4.3. RF Core Designs

The RF core design for each half of the Doherty amplifier is relatively straightforward, especially at 30 MHz. The output matching can be achieved with a simple L

Fig. 4.4 Simulated fundamental currents vs. V\textsubscript{L} and V\textsubscript{PC}.
match to transform $R_L$ to 50 $\Omega$. Selecting $V_{CC} = 3.5$ V and $P_{MAX} = 33$ dBm (2 watts), $R_L$ is calculated as

$$P_{MAX} = \frac{V_L (P_{MAX})^2}{2R_L} = \frac{V_{CC}^2}{2R_L}$$ (48)

$$R_L = \frac{V_{CC}^2}{2P_{MAX}} = \frac{3.5^2}{2 \cdot 2W} = 3.1 \Omega.$$ (49)

Due to the effect of the transistor knee, simulations show a practical value of $R_L = 2.8 \Omega$ is needed to achieve 2 watts output, and the shunt C series L matching network was designed accordingly.

The $V_{CC}$ collector supply was connected to each transistor through a parallel resonant choke, tuned at the fundamental frequency. This provides for the short circuit at all harmonics that is required for class B and class C operation. To prevent significant power loss at the fundamental frequency and to provide a low impedance at the harmonics, a high Q resonator with small values of reactance is required, which will be discussed in more depth in Section 5.1. A second harmonic trap was later added to both amplifiers to provide a better short than was practical with the parallel resonator circuit alone.

It was decided to eliminate the interstage match for this prototype design and directly couple the output current from the first stage to the base of the second stage. Excess gain at 30 MHz allows the mismatch loss of this approach to be ignored, and will certainly be unacceptable at 1900 MHz. But in this case, adequate performance and circuit simplicity were achieved. Note that while simulating and adjusting the final circuit, modifying the Q1HP parallel resonant circuit was found to help overall
performance. The impedance from the Q1_HP collector is now transformed down in impedance somewhat, but the primary benefit seems to be the phase shift this matching circuit provides, better aligning the phases of I_HP and I_LP at the power combining point.

For the input matches, the input impedances of Q1_LP and Q1_HP at full power were transformed to approximately 100 $\Omega$, so the parallel combination would present a low input VSWR to the 50 $\Omega$ source at full power. Simulation predicts better than 1.3:1 input VSWR at 30 MHz.

4.4. Impedance Inverters

Lumped element impedance inverters were used because quarter wave lines are too large to fit in a practical cell phone PA module. The use of 30 MHz for the prototype further forces lumped element inverters because of the ten meter wavelength.

The characteristic impedance of the output impedance inverter, for an $\alpha = 0.25$ is, from (9),

$$Z_{INV} = \frac{R_t}{\alpha} = \frac{2.8}{0.25} = 11.2 \Omega$$

and a $Z_{INV} = 11 \Omega$ design was implemented. The lumped equivalent $\pi$-network values (Fig. 4.2) for an impedance inverter [14] are

$$L_{OUT} = \frac{Z_{INV}}{\omega_0} = \frac{11}{2\pi \cdot 30MHz} = 58.4 nH$$

and

$$C_{OUT} = \frac{1}{Z_{INV} \cdot \omega_0} = \frac{1}{11 \cdot 2\pi \cdot 30MHz} = 482 pF$$
and are implemented in the final design. The input values are calculated similarly, for a $Z_{\text{INV}} = 100 \, \Omega$. Later the values in the input impedance inverter were modified to compensate for a phase error at the power combining point.

Getting the output currents to combine in phase at the combining point (at $V_L$ in Fig. 4.2) was a significant problem in the simulated design and later in the hardware. The phase shift through an amplifier with two saturated stages and dynamic bias levels is a highly nonlinear problem, so it is unrealistic to expect that both amplifiers will have the same phase delay through them, especially as $V_{\text{PC}}$ is swept. It is necessary to characterize each amplifier and modify the input impedance inverter or other matching networks to provide the proper phase delay, and to make sure this balance is maintained over the PAs operating conditions, such as temperature and battery voltage. This was found to be a surmountable problem, both in simulation and in the final hardware.

The final simulation circuit is shown in Fig. 4.5. Note that the offset voltage was implemented as a DC source in series with the bias of $Q_{1\text{HP}}$, but not $Q_{2\text{HP}}$. It was found through simulation experiments that the “stall” effect described earlier could be somewhat reduced by biasing $Q_{2\text{HP}}$ directly from the $V_{\text{PC}}$ supply. $Q_{2\text{HP}}$ is in class C mode near $V_{\text{PC}} = 1.2 \, \text{V}$, so it won’t turn on until $Q_{1\text{HP}}$ has turned on and increased the drive level sufficiently. Future work in still needed to reduce the “stalled” region.

Note also that the Q of the inductors used in the RF output circuitry are critical to efficiency (discussed in Section 5.1). These Qs were measured after the hardware was built, and put back into the simulation to increase its accuracy.
4.5. Simulation Results

Agilent’s Advanced Design System (ADS) was used to simulate and refine the prototype design. Each two-stage amplifier was simulated separately, first with DC and S-parameter simulations to ensure basic circuit integrity, check initial tuning, and check the stability of the two-stage amplifier. Then harmonic balance simulations were used to fine tune each amplifier. Finally, the two amplifiers and the impedance

![Diagram of the final prototype simulation circuit.](image)

Fig. 4.5 Final prototype simulation circuit.
inverters were combined to form the base control Doherty amplifier and the composite performance was simulated and fine tuned. The input and output impedance inverters and the interstage match between $Q_{1\text{HP}}$ and $Q_{2\text{HP}}$ were adjusted to help phase match the two halves of the Doherty amplifier, thus providing maximum power combining.

Fig. 4.6 compares the simulated output stage efficiency to the ideal base control Doherty analysis from Fig. 3.13. The simulation follows the ideal base control Doherty within 15 percentage points throughout the power sweep. The difference between the two curves is attributable to non-ideal behavior in the output transistors and losses in the matching networks. For example, the efficiency of $Q_{2\text{LP}}$ alone
(without the impedance inverter or output match) was measured to be 81% at 21 dBm, suggesting that 10% of the ideal efficiency (91% ideal efficiency minus 81%) was lost in $Q_{2HP}$ and the parallel resonant circuit, and the remaining 5% (from 81% to the 76% peak at 21 dBm in Fig. 4.6) was lost in the impedance inverter and output match.

Fig. 4.7 shows the overall efficiency when the collector currents from the first stages are added. This is essentially equal to the PAE of the amplifier, except when the power out reduces to within 10 dB of the input power (the input power for all simulations is fixed at 0 dBm). The overall efficiency is 69% at 21 dBm and 66% at 33 dBm.

![Graph of total collector efficiency vs. output power](image)

**Fig. 4.7** Simulated efficiency of all stages of the base control Doherty.
As $V_{PC}$ is increased, $Q_{2LP}$ reaches voltage saturation at the transition point. The voltage swing on the $Q_{2LP}$ collector should then remain saturated (its magnitude near $V_{CC}$) throughout the transition region. Fig. 4.8 shows this constant voltage due to the Doherty action (~3.5 V after $V_{PC} = 0.8$ V).

![Graph showing Doherty action](image)

Fig. 4.8 Doherty action: constant voltage saturation in transition region.

Next, simulation results for a single two-stage base controlled amplifier ($Q_{1HP}$ and $Q_{2HP}$) will be compared to the base control analysis in Chapter 3. Fig. 4.9 compares the simulated $Q_{2HP}$ efficiency to the ideal base control and ideal class B PA efficiency. Note that the simulation agrees well with the ideal base control curve until full power is reached, where the non-ideal effects of the transistor and losses in the matching network begin to “fold over” the power curve.
Fig. 4.10 shows the Q2_HP base voltage and current waveforms as V_PC is swept through its full range, with each line representing a different V_PC. V(PC) = 1.6 V is shown in red, and demonstrates the very small conduction angle at this low power level. The shape of these curves agrees well with those predicted in Fig. 3.10b.

Fig. 4.9  Base control PA efficiency, simulation and ideal.

Fig. 4.10  Q2_HP base voltage and current waveforms. Each line represents a different value of V_PC as it is swept from 0.6 to 2.2 V (0.1 V steps).
Fig. 4.11 compares the RF current from the driver stage, Q1\text{HP}, to that of the output stage. In Section 3.4 the assumption was made that the drive levels from the first stage are reasonably constant in the upper power region of the base control amplifier, and the drive level changes rapidly in the lower power region. Fig. 4.11 shows that the drive level (indicated by the Q1\text{HP} output current) does indeed change throughout the V\text{PC} sweep. However, the slope for most of the Q2\text{HP} output current curve is greater than that of the input drive curve, and the simulation results in Fig. 4.9 agree closely with the analysis, so the simplification used earlier appears to be justified. Future work could be done to develop an analysis that includes an increasing drive level from Q1\text{HP}.

Fig. 4.11  Fundamental magnitude of Q1\text{HP} and Q2\text{HP} collector currents. The Q1\text{HP} current represents the drive current to the 2\textsuperscript{nd} stage, and the Q2\text{HP} current represents the output current.
5. PROTOTYPE RESULTS

5.1. Hardware Construction

Photographs of the final hardware are shown in Fig. 5.1 and Fig. 5.2. The HBT arrays were individually die attached and wire bonded into a large gold plated metal package. This allowed the flexibility to easily change the number of transistor cells used for each stage or replace a damaged die. Multiple wires were bonded between each transistor and package pin to reduce the parasitic resistance and inductance due to bond wires. Then the base of the gold package was soldered to a printed circuit board for good heat sinking and grounding.

The remaining circuit components were added discretely, using the package leads to interconnect them. Toroid inductors were used instead of air core or chip inductors to improve the Qs and reduce coupling between them. As mentioned in Section 4.4, the Qs of the output match inductor, output impedance inverter inductor, and output parallel resonant circuits were critical to the efficiency of the circuit. An early version of the hardware used air core inductors with Qs of 69, and the resulting measured efficiency was as much as 20 percentage points lower than in the final circuit, which used inductors with Qs of 110 to 160. The sensitivity to Q can be easily understood by calculating the series or shunt resistances that result from the Q of these inductors, then calculating the amount of power consumed in these resistances while the circuit is in operation. For example, a 26 nH inductor with a Q of 69 was used in the Q2LP parallel resonator in the early hardware. The loss due to this inductor is
Fig. 5.1 Final hardware.

(a) Low power amplifier. (b) High power amplifier.

Fig. 5.2 Close-up of transistors.
equivalent to adding a 340 Ω shunt resistor from the collector of Q2_{LP} to ground. At 21 dBm, the load provided to Q2_{LP} is around 43 Ω, so the 340 Ω resistor consumed 11% of the RF power, obviously having a dramatic effect on efficiency. In the final hardware, this inductor was changed to be 95 nH with a Q of 155. The shunt resistance in this case increases to 2780 Ω and consumes only 1.5% of the RF power.

The impedance inverters also presented a special challenge because of the 2 inch distance between the high power and low power amplifiers. The inductance and resistance of the interconnect wire between the amplifiers turned out to be very significant, especially in the output impedance inverter, since it is operating in a 2.8 Ω impedance environment. The problem was solved by using the interconnect wire as part of the series inductor in the impedance inverter, with the capacitors mounted as close as possible to the amplifiers. Note that very large diameter wires were used for these interconnecting lines, and they were elevated above ground enough to minimize stray capacitance.

5.2. Hardware Results

Fig. 5.3 shows the final efficiency measurements of the hardware circuit. Fig. 5.3a compares the overall efficiency of the base control Doherty amplifier to the simulation results, and Fig. 5.3b compares the efficiency of last stages (using only Q2_{HP} and Q2_{LP} collector currents) to simulation and analysis. The measured efficiency is slightly better at 21 dBm than the simulated efficiency because the Q achieved in some of the inductors was higher than was originally estimated during simulation. The difference between the ideal Q2 efficiency and the measured results
(a) Overall efficiency of measured hardware compared to simulation.

(b) Efficiency of only last stages (Q2LP + Q2HP).

Fig. 5.3 Hardware efficiency results.
is due to the finite Q of the output matching components and the non-ideal transistor characteristics, as discussed in Sections 4.5 and 5.1.

Fig. 5.4 is a collection of voltage waveforms measured with an oscilloscope and low capacitance probe (8pF, 1MΩ). The Q2_{HP} and Q2_{LP} collector voltage waveforms are recorded at 33 dBm and 21 dBm power out. Note that the voltage swing at Q2_{LP} remains saturated (~7 V swing, or 2Vcc) at both 21 dBm and 33 dBm due to the Doherty load modulation effect. Note also that the Q2_{HP} waveform at 33 dBm out is highly non-sinusoidal, possibly because due to non-ideal power combining.
5.3. Lessons Learned

Several new concepts were learned throughout this research. One such concept was that the parasitic effects that commonly degrade a PAs performance at 1900 MHz don’t just disappear at 30 MHz. Parasitic resistance from wiring and inside the circuit elements is always a major problem in a 3 Ω impedance environment, and parasitic inductance can still be significant if the length of interconnecting lines are long enough. An early version of the hardware had to be extensively modified to reduce series resistive, inductive, and even shunt capacitive effects on some of the longer lines.

Another lesson learned was about phase matching between the two sides of a Doherty amplifier and the constraints this places on the design. Whereas in a normal two or three stage amplifier it is possible to tweak the individual tuning elements to get better performance, the Doherty action adds another level of simultaneous conditions that must be met. For instance, adjusting an interstage match component will inevitably change the phase shift through the amplifier, possibly taking the power combining out of phase. In Doherty amplifiers, changes appear to be more easily made in the simulator, where current and voltage waveform measurements are available instantaneously and figures of merit can be calculated more easily than on the bench. Iterating between simulation and hardware measurements seems the best way to optimize a Doherty amplifier.
6. CONCLUSION

A prototype Doherty amplifier for GSM handsets was analyzed, designed and built, using base control techniques to simultaneously control the Doherty action and the output power with a DC control voltage. The concept was demonstrated with hardware at 30 MHz to facilitate analysis and troubleshooting.

A base controlled amplifier was modeled as a class B amplifier at full power, which transitions to class C operation as the output power is reduced, producing superior efficiency at lower power levels. Detailed analysis of Doherty amplifier operation was conducted to better understand how to adapt this technique to constant envelope modulation schemes such as GSM, and maximize the benefits to efficiency at lower power levels.

The base control Doherty amplifier reduces PA current consumption by 34% over a single class B amplifier. Talk time analysis was conducted with new GSM power PDF data and an improvement in talk time of 12% for voice calls and 28% for data calls was estimated.

Areas for future research stemming from this work are:

- Improve the bias offset circuitry so that there is no “stall” in the Doherty action as the power control voltage is ramped.
- Transition the design to the PCS 1900 MHz band and build hardware to evaluate the economic viability of this technique for GSM handsets.
BIBLIOGRAPHY


