Chapter 5
RF Power Amplifier and Linearization Techniques

The radio frequency (RF) power amplifier (PA) is one of the most critical components in designing transmitters in wireless communication systems, and its efficiency dominates the overall efficiency of the transmitter. On one hand, the PA is most power consuming ; for example, in a cellular phone, the battery life is largely determined by the power efficiency of the PA. On the other hand, it is desirable to have the ability to transmit data at the highest possible rate for a given channel bandwidth, i.e., to have high spectral efficiency [1].

5.1 Introduction

There is considerable industrial interest in producing RF PAs with good linearity and power efficiency. These two contradictory requirements can be achieved by using external circuitry to linearize an efficient amplifier [2]. Linearity is the ability of an amplifier to amplify all signals' parts by the same amount, so that all signals are amplified equally and is the most important characteristic of a PA. PAs must be linear to minimize interference and spectral regrowth. However, PAs generally have nonlinear behavior and are basically the main sources of distortion and nonlinearities in the RF transmitter.

Researchers have been focusing on designing more efficient power amplification techniques. It has been shown that the power efficiency of traditional amplification techniques could be improved at a cost of linearity degradation, which may not be tolerable by the standards' requirements. The current state of the art is the design of a moderately linear PA with the additional implementation of a linearization technique [3]. To maximize power efficiency, the amplifier should operate as close to saturation as possible, with the linearization system maximizing the spectral efficiency in this near-saturated region.

In this context, behavioral modeling and linearization of PAs are critical steps in designing high-performance power amplification systems for modern wireless communications infrastructure. Indeed, behavioral modeling of PAs is vital for performance estimation and system level simulation of the transmitter, which provides a time and computationally efficient alternative to the physics based modeling approach [4].

A. Mohammadi et al.: RF Transceiver Design for MIMO Wireless Communications, LNEE 145, pp. 77–128.
springerlink.com

5.2 Transmitter Systems Parameters

The quality of a PA can be characterized by a number of specifications, many of which are provided in detail in the following sections.

5.2.1 Gain

The gain of an amplifier is the ratio of the output amplitude or power to the input amplitude or power, which is usually measured in decibels (dB). The power gain of a PA in dB is given by:

$$G(dB) = 10\log\left(\frac{P_{out}}{P_{in}}\right) \tag{5.1}$$

where P_{in} is the input power and P_{out} is the output power.

5.2.2 Bandwidth

The bandwidth of an amplifier is the range of frequencies for which the amplifier delivers acceptable performance. A well-accepted metric for performance is the half power points (i.e., frequencies where the power gain is half of its peak value). Hence, the bandwidth can be defined as the difference between the lower and upper half power points and is also called a 3-dB bandwidth.

5.2.3 Noise Figure

The noise factor, *F*, is a metric that gives an indication of noise added by the circuit and is defined as the input signal-to-noise ratio (SNR) divided by the output SNR, which is given by:

$$F = \frac{SNR_{in}}{SNR_{out}} \tag{5.2}$$

Since the amplifier always introduces noise, the SNR at the output is always less than that at the input; therefore, the noise factor is always greater than unity.

If the noise factor is expressed in dB, it is called a noise figure (NF) for the circuit, which is obtained as:

$$NF(dB) = 10\log(F) = SNR_{in,dB} - SNR_{out,dB} \tag{5.3}$$

5.2.4 Power Efficiency

Power efficiency in PAs is a metric to quantify the ability of a system to transform the given input power to useful output power. The most power-consuming component of

a wireless transmitter is the PA; and, in fact, the PA is the dominant component in determining the overall power efficiency of a wireless transmitter [5].

There are two forms of input power to the PA: direct current (DC) via the biasing circuits to the gate and drain of the transistor, and RF power at the input of the PA. The RF input power is amplified by transforming the DC power into RF power at the drain level. Part of the DC energy is dissipated as heat, which reduces the PA power efficiency.

There are three different definitions of power efficiency in the literature, which are the total efficiency, the drain efficiency, and the power-added efficiency [6].

5.2.4.1 Total Efficiency

The total efficiency can be obtained as the ratio of the RF power delivered at the output of the PA to the summation of the RF and DC powers that entered the PA. The total efficiency is obtained by:

$$\eta_t = \frac{P_{out}}{P_{dc} + P_{in}} \tag{5.4}$$

where η_t is the total efficiency of the PA; and, P_{dc}, P_{in}, and P_{out} are the DC and RF powers at the input and output of the PA, respectively.

5.2.4.2 Drain Efficiency

Drain efficiency, $\eta_{D\,trans.}$, is defined as the ratio of the RF power to the DC power at the drain level, which is specific to field-effect transistors (FETs). This efficiency is called collector efficiency in bipolar junction transistors (BJTs). Drain efficiency is given by:

$$\eta_{D\,trans.} = \frac{P_{RF\ Drain}}{P_{dc\ Drain}} \tag{5.5}$$

where $P_{dc\ Drain}$ and $P_{RF\ Drain}$ are the DC and RF powers, respectively, at the drain level of the transistor.

The DC power at the gate level of a transistor is usually very small compared to the DC power at the drain level. Hence, the DC power at the drain level of the transistor, $P_{dc\ Drain}$, can be approximated as the total DC power entering the PA, P_{dc}. By considering losses in the matching networks, the drain efficiency, η_D, is given as the ratio of the RF output power, P_{out}, to the DC power that entered the PA, P_{dc}, which is expressed as:

$$\eta_D = \frac{P_{out}}{P_{dc}} \tag{5.6}$$

5.2.4.3 Power-Added Efficiency

The power-added efficiency (PAE), η_{PAE}, demonstrates the ability of the PA to transform DC power to RF power. PAE can be obtained as the ratio of the added power, which is the difference between the output RF power, P_{out}, and the input RF power, P_{in}, to the DC power, P_{dc}. PAE is given by:

$$\eta_{PAE} = \frac{P_{out} - P_{in}}{P_{dc}} \tag{5.7}$$

5.2.5 P1dB

The 1 dB compression point (P1dB) is a measure of amplitude linearity. The gain of an amplifier compresses when the output signal level enters the compression region before it reaches saturation. Higher output power corresponds to a higher compression point. It is an input (or output) power for which the gain of the PA is 1 dB less than the small-signal gain (ideal linear gain).

The maximum saturation point corresponds to the point where the PA reaches its output maximum power. This maximum power is called the saturation power $P_{sat[\max]}$. The 3 dB saturation power $P_{sat[dB]}$ corresponds to the power for which the gain of the PA is 3 dB less than the small-signal gain. Figure 5.1 illustrates the P1dB, maximum saturation and 3 dB saturation points for a typical PA. To avoid intermodulation (IM) problems and distortion, the output power needs to be reduced below the P1dB. Manufacturers usually back off about 10 dB from the P1dB point.

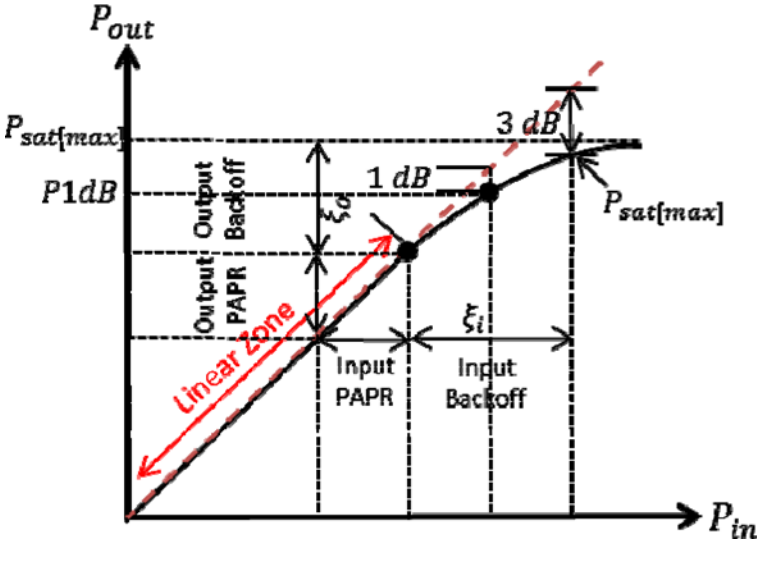

Fig. 5.1 P1dB and saturation power definition for PAs

5.2.6 IP3

A third-order intercept point (IP3) is a widely used metric in PAs, which gives information about the linearity of an amplifier. A higher IP3 means better linearity and lower distortion generation. It is a theoretical point at which the desired output signal and the third-order IM (undesired) signal are equal in levels considering an ideal linear gain for the PA. The theoretical input point is the input IP3 (IIP3) and the output power is the output IP3 (OIP3).

The two-tone test, as shown in Figure 5.2, is commonly used to determine IP3, which uses two closely spaced frequencies. When two signals at frequencies f_1 and f_2 are incident on an amplifier, the output of the amplifier contains these two signals, as well as IM products at frequencies $mf_1 \pm nf_2$, where $m+n$ is known as the order of the IM product. The ratio of third-order IM products and the carrier is known as third-order IM (IM3). IM3 products are important since their frequencies, $2f_1 - f_2$ and $2f_2 - f_1$, fall close to the desired signal, which makes filtering of IM3 an issue. IM3 for $2f_1 - f_2$ is given by:

$$IM3(dBc) = 10\log\left(\frac{P_{2f_1-f_2}}{P_{f_1}}\right) \tag{5.8}$$

and IM3 for $2f_2 - f_1$ is expressed by:

$$IM3(dBc) = 10\log\left(\frac{P_{2f_2-f_1}}{P_{f_2}}\right) \tag{5.9}$$

where p_{f_1}, p_{f_2}, $p_{2f_1-f_2}$, and $p_{2f_2-f_1}$ are the power outputs at frequencies f_1, f_2, $2f_1 - f_2$, and $2f_2 - f_1$, respectively. IM3 is measured in units of dBc, because it is calculated relative to the main tone power output.

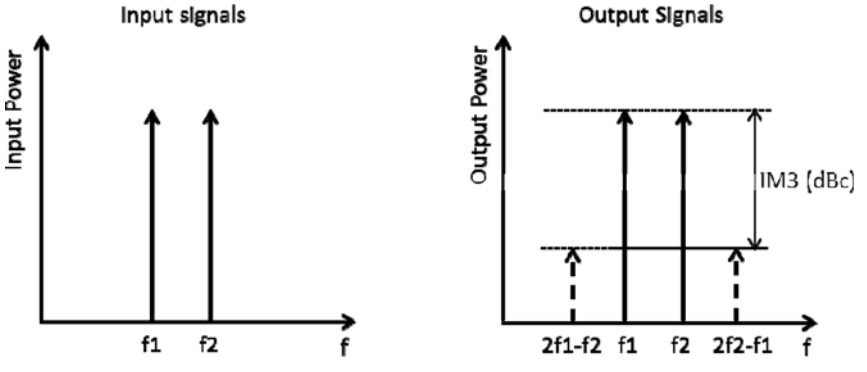

Fig. 5.2 Third-order intermodulation distortion (IM3)

5.2.7 PAPR

A metric used to quantify required linearity is called the peak-to-average power ratio (PAPR), which is the ratio between the peak power P_{peak} (related to peak amplitude) and the average power P_{avg} (related to mean amplitude) of a signal. It is also called the crest factor and is given by:

$$PAR(dB) = 10\log\left(\frac{\max(|x(t)|^2)}{mean(|x(t)|^2)}\right) = 10\log\left(\frac{P_{peak}}{P_{avg}}\right) \tag{5.10}$$

5.2.8 Power Back-Off

An amplifier appears linear for sufficiently small departures from its bias conditions. The power back-off is defined as the ratio between the PA's saturation power to the RF signal's mean power. There are two types of power back-off: input power back-off and output power back-off. The back-off at the input of the PA, ξ_i, is obtained by:

$$\xi_i = P_{i,sat} - P_{i,mean} \quad (dB) \tag{5.11}$$

where $P_{i,sat}$ and $P_{i,mean}$ are the saturation power and mean signal power at the input of the PA, respectively.

Similarly, the back-off at the output of the PA, ξ_o, is given by:

$$\xi_o = P_{o,sat} - P_{o,mean} \quad (dB) \tag{5.12}$$

where $P_{o,sat}$ and $P_{o,mean}$ are the saturation power and mean signal power at the output of the PA, respectively.

The input power back-off and output power back-off are illustrated in Figure 5.1.

5.2.9 ACPR

The adjacent channel power ratio (ACPR) is a critical figure of merit in the evaluation of the IM distortion performance of RF PAs. It is a measure of spectral regrowth and appears in the signal sidebands. ACPR is defined as the ratio of power in a bandwidth adjacent to the main channel to the power within the main signal bandwidth.

The ACPR for the right side of the power spectral density (PSD) can be defined as:

$$ACPR(Right) = \frac{\int_{f_c+\Delta f-\frac{B}{2}}^{f_c+\Delta f+\frac{B}{2}} PSD(f)df}{\int_{f_c-\frac{B}{2}}^{f_c+\frac{B}{2}} PSD(f)df} \tag{5.13}$$

Similarly, the ACPR for the left side of the PSD can be obtained as:

$$ACPR(Left) = \frac{\int_{f_c-\Delta f-\frac{B}{2}}^{f_c-\Delta f+\frac{B}{2}} PSD(f)df}{\int_{f_c-\frac{B}{2}}^{f_c+\frac{B}{2}} PSD(f)df} \tag{5.14}$$

where f_c is the carrier frequency, B is the bandwidth of the modulated signal, and $PSD(f)$ is the power spectral density at frequency f. Figure 5.3 illustrates the way that ACPR is calculated.

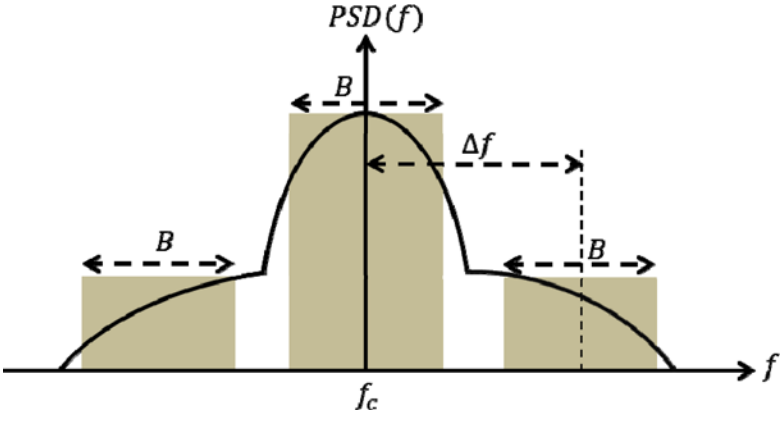

Fig. 5.3 Adjacent channel power ratio (ACPR)

5.2.10 EVM

The effect of noise, distortion, and in-phase / quadrature (I/Q) imbalance on signal quality can be analyzed based on its error vector magnitude (EVM) measurement. The EVM is a common metric for the fidelity of the symbol constellation. Ideally, a signal sent by a transmitter or received by a receiver would have all constellation points precisely at the ideal locations. However, several deficiencies in implementation, such as I/Q imbalance, noise and carrier leakage, cause the deviation of constellation points from their ideal locations. In other words, EVM is a measure of how far the actual (measured) points are from the ideal locations. The discrepancy between actual and ideal symbol vectors is quantified through an error vector, as shown in Figure 5.4.

The EVM is the ratio of the power of the error vector to the power of the reference vector related to the ideal constellation. The EVM can be defined in decibels (dB) or percentage (%) as:

$$EVM\,(dB) = 10\log_{10}\left(\frac{P_{error}}{P_{ideal}}\right) = 10\log_{10}\left(mean\left(\frac{(I_{actual}-I_{ideal})^2+(Q_{actual}-Q_{ideal})^2}{I_{ideal}^2+Q_{ideal}^2}\right)\right) \tag{5.15}$$

$$EVM\,(\%) = \sqrt{\frac{P_{error}}{P_{ideal}}}\times 100 \tag{5.16}$$

where P_{error} and P_{ideal} are the root mean square (RMS) power of the error vector and the outermost (highest power) point in the reference signal constellation, I_{ideal} and Q_{ideal} are the ideal in-phase and quadrature signals, and I_{actual} and Q_{actual} are the transmitted in-phase and quadrature signals.

EVM can be also represented as:

$$EVM = \sqrt{\frac{\frac{1}{N}\sum_{n=1}^{N}\left|S_{ideal,n}-S_{actual,n}\right|^2}{\frac{1}{N}\sum_{n=1}^{N}\left|S_{ideal,n}\right|^2}} \tag{5.17}$$

where N is the number of unique symbols in the constellation.

EVM is useful in quantifying the amount of noise and distortion in a transmitter. The random distribution of the error vectors is caused by noise, while patterns in the error vector measurements correlated with symbol amplitude and phase may be caused by distortion.

It is common to characterize AM/AM (amplitude modulation to amplitude modulation) distortion and AM/PM (amplitude modulation to phase modulation) distortion, which is shown in Figure 5.5. Variation in the transmitter gain causes AM/AM distortion. Phase shifting of the carrier correlated to the signal amplitude introduces AM/PM distortion [7].

Undesired expansion of the output spectrum is also caused by distortion. Nonlinearity of a PA that can be quantified as AM/AM and AM/PM distortion generates counterfeit energy near the fundamental and harmonics of the carrier frequency. This broadening in the output spectrum is referred to as spectral regrowth or out-of-band distortion, which deteriorates the ACPR. The transmitter output spectrum showing spectral regrowth is illustrated in Figure 5.6. Spectral regrowth is a potentially significant problem and demands special attention during the transmitter design process. The nonlinearity in the device can be compensated for using linearization techniques [4], which are covered in Section 5.4.

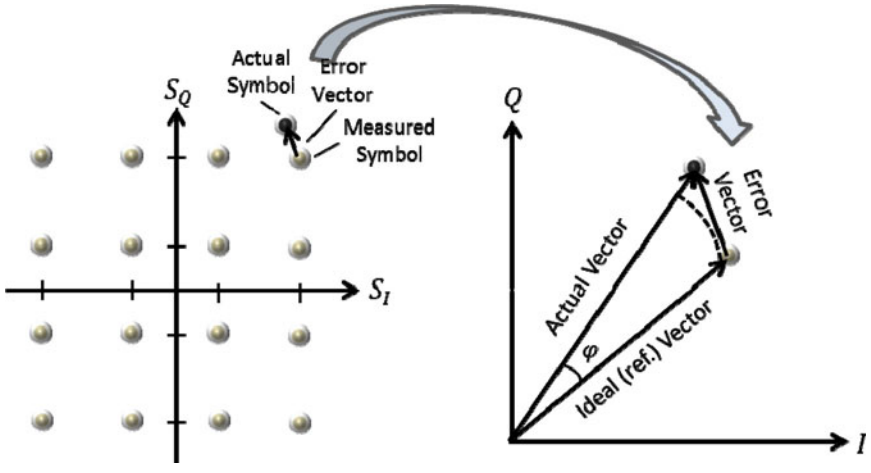

Fig. 5.4 Error vector is the difference between the actual and ideal symbol vectors

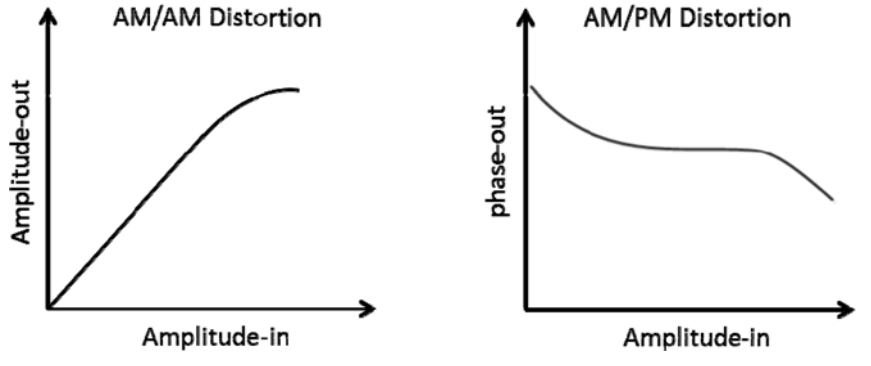

Fig. 5.5 AM/AM distortion and AM/PM distortion in transmitters

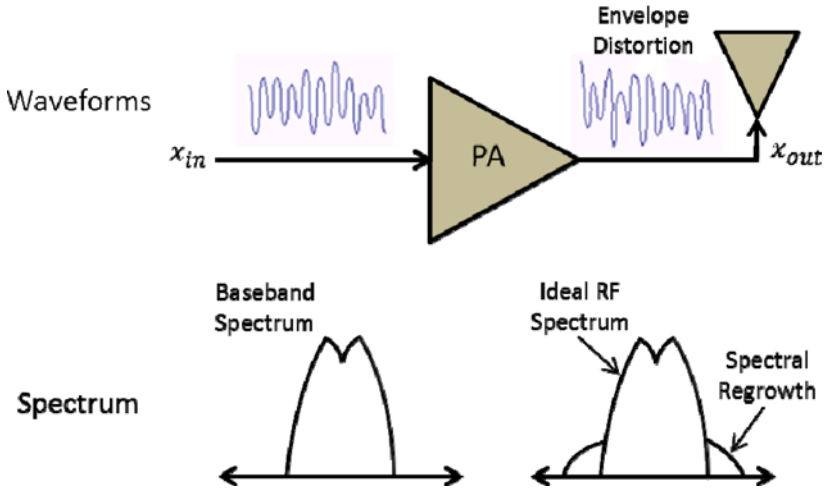

Fig. 5.6 Spectral regrowth in transmitter output spectrum

5.2.11 Memory Effect

In a two- or multi-tone IM test, if the amplitude or phase of the IM signals is affected by the tone frequency spacing, the PA exposes memory effects. Memory effects can be explained as time lags between the AM/AM and AM/PM responses of the PA. Electrical and thermal are two types of memory effects.

The electrical memory effect is produced by poor gate and drain decoupling in FET and base and collector decoupling in BJT at low frequencies. Electrical memory effects cause the distortion of the envelope currents and result in IM asymmetry. The memory effect is more significant for class AB PAs than class A, with a reduced conduction angle.

Electrical memory effects are generated by non-constant node impedances within frequency bands. Most of them are produced by frequency-dependent envelope impedances. The thermal memory effect is generated by the junction temperature that is modulated by the applied signal envelope.

5.3 RF Power Amplifiers

The wireless and satellite communications communities are always searching for radios that are more power efficient. However, the power efficiency is greatly dominated by the efficiency of the RF PA in the transmitting path. For a transistor, the output power, gain and power efficiency depend on the biasing and matching conditions, which define the PA's class of operation. A diagram showing the conduction angles, as well as the input drive levels, is presented in Figure 5.7 for the different classes of power amplifiers: (i) continuously driven class A, AB, B and C PAs, (ii) saturated PAs and (iii) class D, E and F switching-modes PAs [1].

By reducing the conduction angle, the efficiency of the PAs is increased. The efficiency is even higher when operating the PA in a switching mode; however, efficiency enhancement is accompanied with linearity degradation. The higher the efficiency when passing from one class to another, the poorer is the linearity performance.

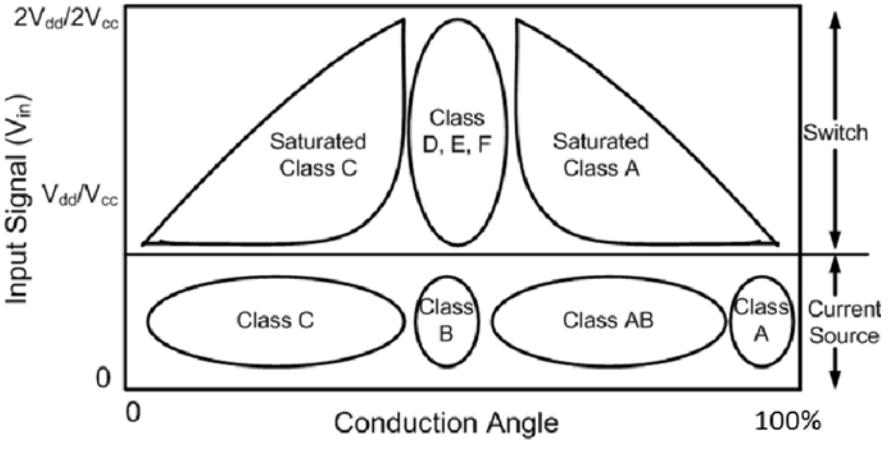

Fig. 5.7 Classes of power amplifiers [1]

Time-varying envelope signals are implemented in most communication systems that have recently been deployed or are to be deployed in the near future. PAs are required to operate in their back-off region with such signals to meet the required linearity. This linearity is defined either by the ACPR or EVM in the context of wireless applications. The back-off level setting is generally a function of the input signal's PAPR. Unfavorably, the power efficiency of the PA decreases as the back-off level increases. Therefore, this leads to the design of very low efficiency amplifiers that need large DC power modules for fixed terminals or shorten battery life for mobile terminals. In fact, linearization techniques are being used to improve achievable power efficiency [1].

5.3.1 Linear PAs (Classes A, AB, B, C)

This section discusses several types of classic RF PAs, which are called classes A, AB, B and C. Power amplifiers, depending on the class of operation, behave with different levels of nonlinearity and provide different levels of efficiency. These PAs, in terms of levels of linearity and power efficiency, are described; and, the tradeoff between the linearity and efficiency is discussed.

A general block diagram of a PA is shown in Figure 5.8, which consists of a transistor with its biasing circuits and the input and matching networks. The required current and voltage to bias the transistors are provided by the biasing circuits. The input and output matching networks are designed to transfer the maximum power through the PA.

The class of operation of the PA relies upon the biasing level of the transistor. For instance, the DC current consumption variation of the FET transistor is shown in Figure 5.9, as a function of the gate voltage for a given drain voltage. The DC current in this figure cuts into two regions. In the first region, the current is zero; therefore, the transistor is off. In the second region, the current changes almost linearly with the gate voltage. The gate voltage at the edge of these two regions is described as the pinch-off voltage. The gate bias regions corresponding to class A, AB, B and C PAs are also illustrated in Figure 5.9.

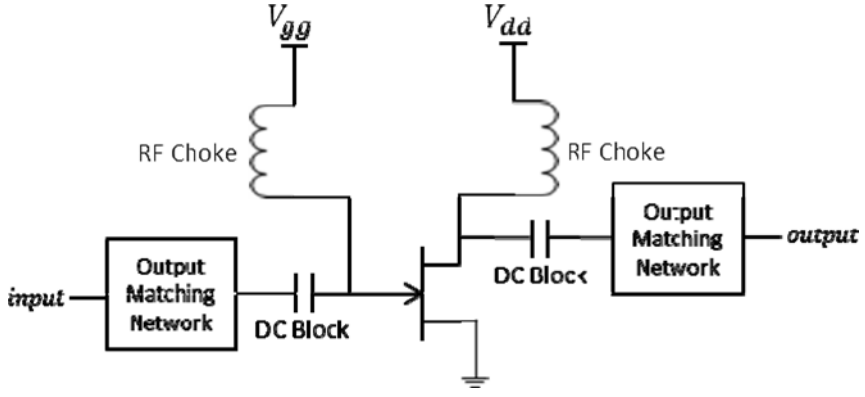

Fig. 5.8 A general block diagram of a PA

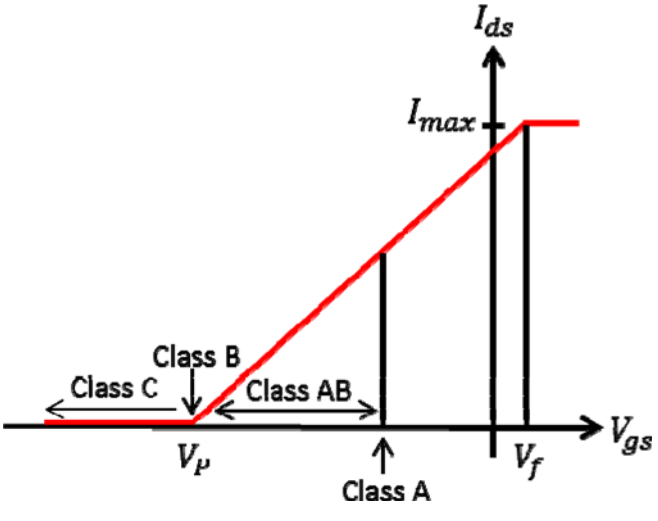

Fig. 5.9 The DC current consumption versus the gate voltage of the FET transistor

In class B operation, the transistor starts conducting as soon as the input RF signal is presented at the input of the PA; in fact, the PA consumes power as long as the RF signal is present at the input. Operation of the PA in classes A, AB and C can be achieved by biasing the transistor at the gate level, below or above the pinch-off voltage. In class A, the transistor is on, even when there is no input RF signal. Hence, the transistor conducts the input RF signal with no limitation on the input power level up to the maximum level that the transistor can handle. In contrast, in class C, the transistor is off as long as the input RF signal amplitude does not exceed a power threshold. Class A PAs have a linear power transfer function, while it becomes nonlinear moving toward class C.

The conduction angle, α, is defined as the angular period of the input RF signal cycle in which the transistor conducts. It varies between 0 to 2π, and $\alpha = \pi$ corresponds to class B operation. The conduction angle is calculated by:

$$\alpha = 2\cos^{-1}\left(1 - \frac{V_P}{V_{gs}}\right) \tag{5.18}$$

where V_P is the pinch-off voltage, and V_{gs} is the gate-source voltage.

It was theoretically proven [8] that the maximum drain efficiency of the PA can be related to the conduction angle by the following equation:

$$\eta_{\max} = \frac{\alpha - \sin(\alpha)}{4\sin\left(\frac{\alpha}{2}\right) - 2\alpha\cos\left(\frac{\alpha}{2}\right)} \tag{5.19}$$

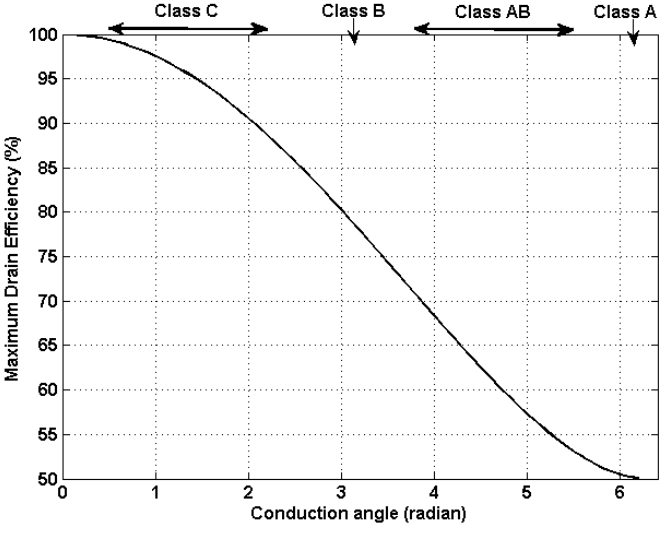

Fig. 5.10 The maximum drain efficiency versus the conduction angle

Based on the previous equation, the maximum drain efficiency of the PA increases from 50% for a class A PA with a conduction angle of 2π to 100% for a class C PA with a conduction angle of 0. This is achieved by decreasing the conduction angle. The theoretical maximum drain efficiency of the PA for a conduction angle between 0 to 2π is demonstrated in Figure 5.10. It can be seen that the maximum efficiency of the PA increases from class A to class C, where the class A PA has a linear power transfer function and the class C PA has a nonlinear power transfer function. Indeed, there should be a tradeoff between the efficiency and linearity of the PA.

Considering the preceding discussion, the choice of a proper class of operation for a PA in wireless transmitters is fairly challenging. Signal quality degradation can be avoided by highly linear PAs, but it results in poor power efficiency. The overall linearity and efficiency of wireless transmitters are highly affected by the PA, since it is the dominant source of nonlinearities and power consumption. A class AB PA supplies an acceptable tradeoff between linearity and power consumption.

5.3.1.1 Class A PA

The class A PA has the highest linearity over the other classes of operation. In the class A amplifier, the conduction angle is set to $\alpha = 360^\circ$; and, current flows constantly and is not cut off during any part of the cycle, acting as a current source. To achieve high linearity and gain, the amplifier's base and drain DC voltage has to be chosen properly so that the amplifier can operate in a linear region. There is continuous loss of power in the amplifier, since it is constantly carrying current. The class of operation A is distinguished owing to the fact that it is, in theory, perfectly linear, but inefficient.

In a Class A PA, the gate and drain of a FET transistor are biased according to:

$$V_{gg} = \frac{V_P + V_f}{2} \tag{5.20}$$

$$V_{dd} = \frac{V_{\max} + V_{knee}}{2} \tag{5.21}$$

where V_P is the gate pinch-off voltage, V_f is the gate diode forward voltage, $V_{\max}$ is the maximum tolerable drain voltage of the transistor, and V_{knee} is the drain knee voltage. V_P, V_f, $V_{\max}$, and V_{knee} are inherent parameters of a FET transistor and can be collected from the device data sheet.

To achieve the maximum drain current swing without clipping, the maximum amplitude of the RF gate voltage should be chosen as $V_f - V_{gg}$, which is illustrated in Figure 5.11 (a). The input power corresponding to this gate voltage is called the maximum linear input power. Consequently, the device conducts for the entire RF cycle. The maximum drain current swing, and thus maximum linear output power, can be obtained by proper selection of the load impedance. The gate and drain voltages and drain current waveforms for a class A PA are illustrated in Figure 5.11, as a function of the angular phase, θ.

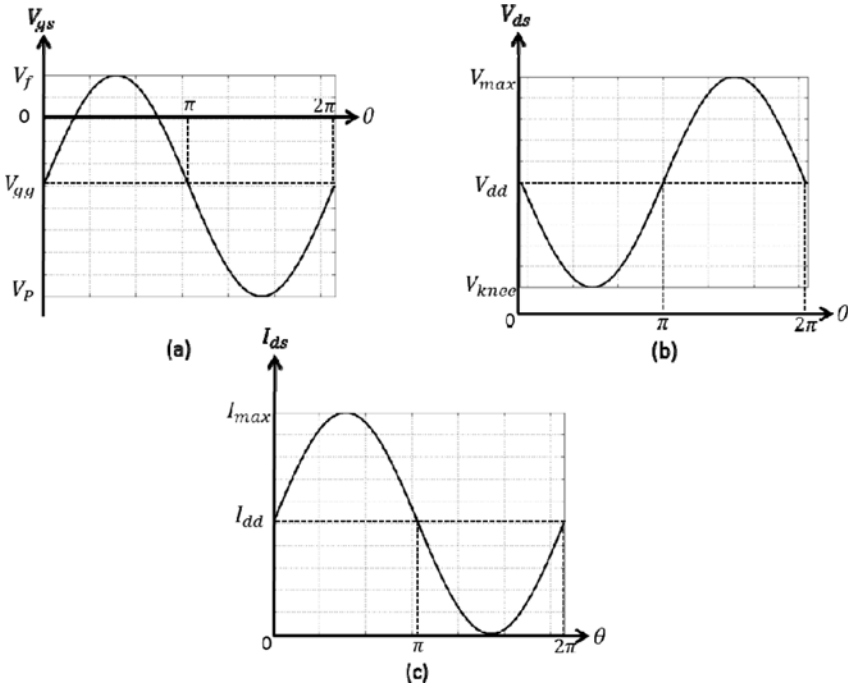

Fig. 5.11 Class A: (a) gate voltage, (b) drain voltage and (c) drain current waveforms

The drain current shown in Figure 5.11 (c) is the response of the transistor, which is assumed to be an ideal voltage-dependent current source, to the applied gate voltage shown in Figure 5.11 (a). The drain voltage is also represented in Figure 5.11 (b) with the assumption of providing the optimal load for the maximum linear output power. The gate and drain voltages and current corresponding to the waveforms in Figure 5.11 can be described as:

$$V_{gs}(\theta) = V_{gg} + v_g \cdot \sin\theta \text{, where } v_g = V_f - V_{gg} \tag{5.22}$$

$$V_{ds}(\theta) = V_{dd} - v_d \cdot \sin\theta \text{, where } v_d = V_{dd} - V_{knee} \tag{5.23}$$

$$I_{ds}(\theta) = I_{dd} + i_d \cdot \sin\theta \text{, where } i_d = I_{dd} = I_{\max} / 2 \tag{5.24}$$

Thus, the optimal load, $R^A_{opt.}$, for the maximum drain current swing can be calculated as:

$$R^A_{opt.} = \frac{v_d}{i_d} = \frac{V_{dd} - V_{knee}}{I_{\max} / 2} = 2\frac{V_{dd} - V_{knee}}{I_{\max}} \tag{5.25}$$

The maximum output power at peak drain efficiency is obtained if the transistor works on a load line of $R^A_{opt.}$. The required DC power, $P^A_{dc,\max}$, maximum RF output power, $P^A_{RF,\max}$, and peak drain efficiency, $\eta^A_{D,\max}$, for class A operation can be obtained using the following equations:

$$P^A_{dc,\max} = V_{dd} I_{dd} = \frac{V_{dd} \cdot I_{\max}}{2} \tag{5.26}$$

$$P^A_{RF,\max} = \frac{v_d i_d}{2} = \frac{(V_{dd} - V_{knee}) \cdot I_{\max}}{4} \tag{5.27}$$

$$\eta^A_{D,\max} = \frac{P^A_{RF,\max}}{P^A_{dc,\max}} = \frac{(V_{dd} - V_{knee})}{2V_{dd}} \tag{5.28}$$

if $V_{knee} = 0$, then $\eta^A_{D,\max} = \frac{1}{2}$.

The dissipated power in the transistor, which is the difference between the injected DC power to the amplifier and the RF output power produced by the PA, can be obtained as:

$$P^A_{diss} = \frac{1}{2\pi} \int_0^{2\pi} V_{ds}(\theta) \cdot I_{ds}(\theta) d\theta = P_{dc} - P_{RF} = \frac{V_{\max} \cdot I_{\max}}{8} \tag{5.29}$$

5.3.1.2 Class B PA

As shown in Figure 5.9, in the class B PA, the gate of the transistor is biased at the pinch-off voltage, and the conduction angle is set to $\alpha = 180^\circ$. Hence, the transistor remains in the active region during half of the RF cycle. A class B PA operates at zero quiescent current. This half-sine waveform of the drain current reduces the corresponding DC power; therefore, the power efficiency is increased.

In the class B mode of operation, the gate and drain of a FET transistor are biased using:

$$V_{gg} = V_P \tag{5.30}$$

$$V_{dd} = \frac{V_{\max} + V_{knee}}{2} \tag{5.31}$$

The gate and drain voltages and drain current waveforms for a class B PA are depicted in Figure 5.12. As in the class A mode of operation, the load-line impedance at the fundamental frequency is chosen to obtain the maximum drain voltage and current swing. The harmonics of the fundamental frequency are ideally shortened to ground at the output of the PA.

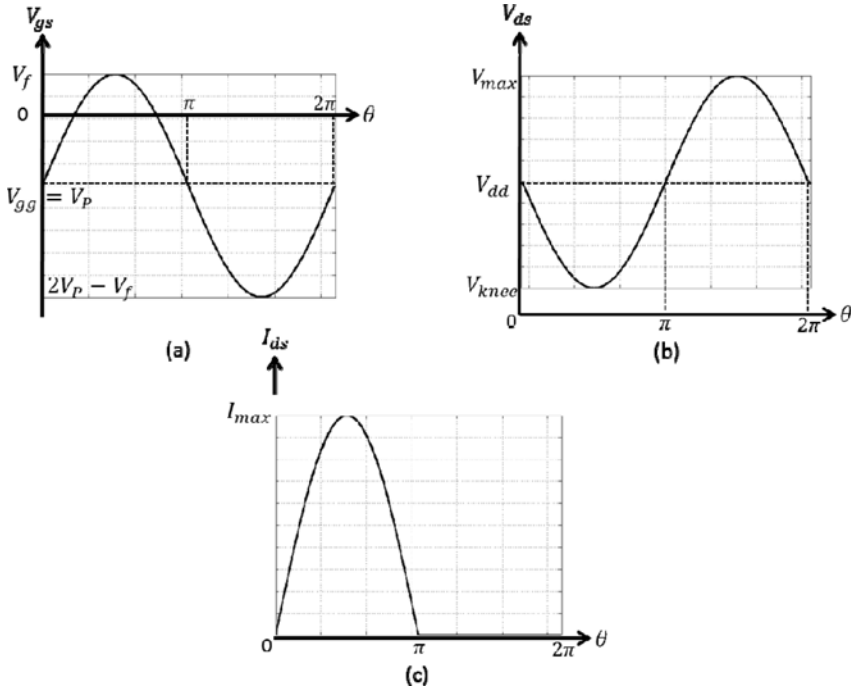

Fig. 5.12 Class B: (a) gate voltage, (b) drain voltage and (c) drain current waveforms

In the class B mode of operation, injecting a sinusoidal waveform to a transistor with the gate biased at the pinch-off voltage results in a half-sine waveform of the drain current that consists of fundamental and even harmonic frequencies. The drain voltage only includes the fundamental frequency, since all higher harmonic voltage components are suppressed by proper harmonic termination. The gate and drain voltage equations are exactly the same as those in (5.22) and (5.23), respectively, but with different V_{gg}. The drain current is calculated as:

$$I_{ds}(\theta) = \begin{cases} I_{\max} \cdot \sin\theta & 0 \le \theta \le \pi \\ 0 & \pi < \theta \le 2\pi \end{cases} \tag{5.32}$$

The amplitude of the DC and the fundamental frequency content of the drain current of a class B PA can be calculated with the following equations:

$$I_{dd} = \frac{1}{2\pi}\int_0^{2\pi} I_{ds}(\theta)d\theta = \frac{I_{\max}}{\pi} \tag{5.33}$$

$$i_{d,1} = \frac{1}{\pi}\int_0^{2\pi} I_{ds}(\theta)\cdot\sin(\theta)d\theta = \frac{I_{\max}}{2} \tag{5.34}$$

Assuming the same maximum drain voltage and current as in the class A mode of operation can be obtained, the peak output power of class B is equal to that of the class A PA, but at the cost of 6 dB more input power requirement. This results in a gain that is lower by 6 dB. The reason is that only half of the input RF signal is amplified by the PA. However, class B mode of operation achieves higher peak efficiency than class A, due to the lower average DC current, and, hence, the lower dissipated DC power.

A class B PA yields maximum power when offered the same optimal load line as that of a class A PA, which is calculated in (5.25). The required DC power, $P^B_{dc,\max}$, maximum fundamental RF output power, $P^A_{RF,\max}$, and maximum power efficiency, $\eta^A_{D,\max}$, for the class B operation mode can be calculated based on the following equations:

$$P^B_{dc,\max} = V_{dd} I_{dd} = \frac{V_{dd} \cdot I_{\max}}{\pi} \tag{5.35}$$

$$P^B_{RF,f_0,\max} = \frac{v_{d,1} i_{d,1}}{2} = \frac{(V_{dd} - V_{knee}) \cdot I_{\max}}{4} \tag{5.36}$$

$$\eta^A_{D,\max} = \frac{P^B_{RF,f_0,\max}}{P^B_{dc,\max}} = \frac{\pi (V_{dd} - V_{knee})}{4 V_{dd}} \tag{5.37}$$

if $V_{knee} = 0$, then $\eta^B_{D,\max} = \frac{\pi}{4} = 78.5\%$

The theoretical optimal efficiency of the class B PA is $\eta^B_{D,\max} = 78.5\%$ higher than that of the class A PA $\eta^A_{D,\max} = 50\%$, with the cost of reduction in the linearity of the device.

The dissipated power in the transistor for a class B mode of operation can be calculated as:

$$P^B_{diss} = \frac{1}{2\pi} \int_0^{2\pi} V_{ds}(\theta) \cdot I_{ds}(\theta) d\theta = P_{dc} - P_{RF} = \frac{4 - \pi}{8\pi} V_{\max} \cdot I_{\max} \tag{5.38}$$

In comparison to (5.29), the dissipated power in the transistor in class B PA is around 27% of the power dissipation in class A PA for the same maximum output power.

In conclusion, a class B PA yields the same maximum output power as class A PA, while producing higher efficiency, but 6 dB less gain. Furthermore, a more complex load network is required, due to the requirement of proper harmonic frequency termination.

5.3.1.3 Class AB PA

A class AB PA compromises between class A and class B operation in terms of efficiency and linearity. The conduction angle is $180^\circ < \alpha < 360^\circ$, so that the transistor remains in the active region for more than half of but less than the full cycle of the input RF signal. Indeed, the efficiency in the class AB mode of operation is between 50% and 78.5%. Class AB PAs offer a wider dynamic range than either class A or B PAs. The linearity of class AB is quite similar to a class A PA before its input power becomes large enough to cause the drain current clipping from the bottom, which is the time when nonlinearity starts and the gain decreases.

5.3.1.4 Class C PA

In the class C mode of operation, the gate bias point is chosen below the pinch-off voltage of the device. In a class C PA, the conduction angle is $\alpha < 180^\circ$, and the transistor remains in the active region for less than half of the RF cycle. Hence, in order to achieve the maximum transistor drain current, $I_{\max}$, more input power than in the class B is needed. The class C PA has higher efficiency than classes A, AB and B, but it is highly nonlinear. It is biased so that the output current is zero for more than one half of an input sinusoidal signal cycle.

In the class C operation mode, the gate and drain of a FET transistor are biased according to:

$$V_{gg} < V_P \tag{5.39}$$

$$V_{dd} = \frac{V_{\max} + V_{knee}}{2} \tag{5.40}$$

Figure 5.13 shows the gate and drain voltages and drain current waveforms for a class C PA. The gate and drain voltage equations are the same as those in (5.22) and (5.23), respectively, but with a different V_{gg}. The conduction angle in Figure 5.13 (c) can be calculated by:

$$V_P = V_{gg} + v_g \cdot \sin\gamma \tag{5.41}$$

where, $\gamma = \frac{\pi}{2} - \theta_0$, then,

$$\theta_0 = \cos^{-1}\left(\frac{V_P - V_{gg}}{v_g}\right) \tag{5.42}$$

Assuming an ideal linear relationship between the drain current, I_{ds}, and the gate voltage, V_{gs}, as demonstrated in Figure 5.9, the drain current can be specified as:

$$I_{ds}(\theta) = \frac{I_{\max}}{V_f - V_P} V_{gs}(\theta) - \frac{I_{\max} \cdot V_P}{V_f - V_P} \tag{5.43}$$

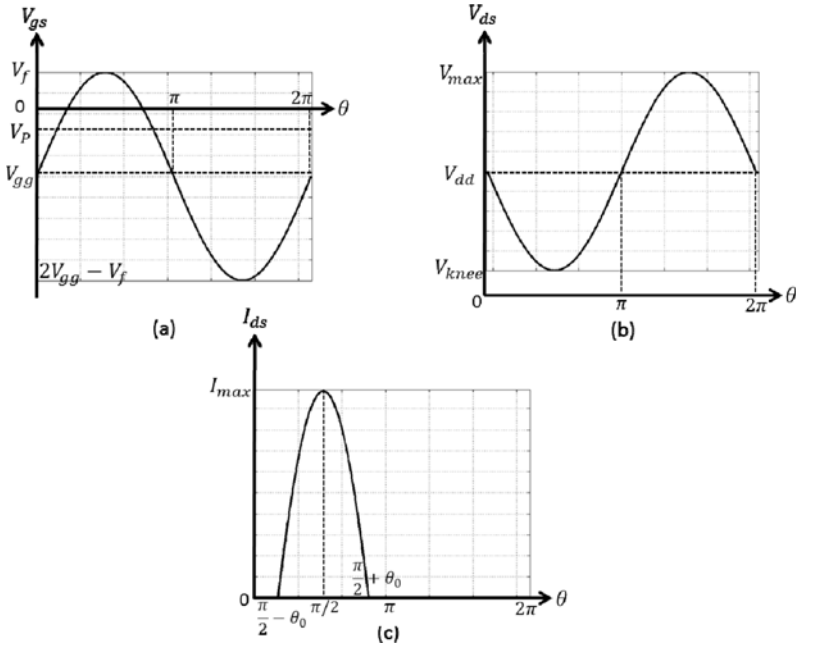

Fig. 5.13 Class C: (a) gate voltage, (b) drain voltage and (c) drain current waveforms

Based on the class C conduction angle and applying (5.22) to (5.43), the drain current can be calculated as:

$$I_{ds}(\theta) = \begin{cases} \dfrac{I_{\max} v_g}{V_f - V_P} \sin\theta + \dfrac{I_{\max}(V_{gg} - V_P)}{V_f - V_P} & \dfrac{\pi}{2} - \theta_0 \le \theta \le \dfrac{\pi}{2} + \theta_0 \\ 0 & otherwise \end{cases} \tag{5.44}$$

This equation is valid for all the classes of operation, if v_g, V_{gg} and θ_0 are suitably chosen. In the class C PA, the drain current waveform includes both odd and even harmonic frequency content. Hence, all higher frequency components need to be short-circuited at the output to have zero drain voltage at all harmonics except at the fundamental frequency.

Using (5.44), the drain current at DC and the fundamental frequency can be derived in a few steps with the following equations:

$$I_{dd} = \frac{1}{2\pi}\int_0^{2\pi} I_{ds}(\theta)d\theta = \frac{I_{\max} \cdot v_g}{\pi(V_f - V_p)}(\sin\theta_0 - \theta_0\cos\theta_0) \tag{5.45}$$

$$i_{d,1} = \frac{1}{2\pi}\int_0^{2\pi} I_{ds}(\theta) \cdot \sin(\theta)d\theta = \frac{I_{\max} \cdot v_g}{\pi(V_f - V_p)}(\theta_0 - \sin\theta_0\cos\theta_0) \tag{5.46}$$

To achieve peak drain current and, accordingly, maximum output power, the required RF gate voltage amplitude of a class C PA should be selected as:

$$v_g = V_f - V_{gg} \tag{5.47}$$

Using (5.41) and (5.47), v_g can be calculated as a function of θ_0 as:

$$v_g = \frac{V_f - V_P}{1-\cos\theta_0} \tag{5.48}$$

By substituting (5.48) in (5.45) and (5.46), the following equations can be obtained:

$$I_{dd} = \frac{I_{\max}}{\pi}\frac{(\sin\theta_0 - \theta_0\cos\theta_0)}{1-\cos\theta_0} \tag{5.49}$$

$$i_{d,1} = \frac{I_{\max}}{\pi}\frac{(\theta_0 - \sin\theta_0\cos\theta_0)}{1-\cos\theta_0} \tag{5.50}$$

The class C optimal load line, $R^C_{opt.}$, for the maximum drain voltage swing is calculated as:

$$R^C_{opt.} = \frac{v^C_{d,1}}{i^C_{d,1}} = \frac{\pi(V_{dd} - V_{knee})}{I_{\max}} \cdot \frac{1-\cos\theta_0}{(\theta_0 - \sin\theta_0 \cdot \cos\theta_0)} \tag{5.51}$$

Biasing the gate deeper in class C decreases the drain current; hence, a larger load resistor is needed to achieve the maximum voltage swing.

The required DC power, $P^C_{dc,\max}$, maximum fundamental RF output power, $P^C_{RF,\max}$, and maximum drain power efficiency, $\eta^C_{D,\max}$, for the class C operation mode can be calculated based on the following equations:

$$P^C_{dc,\max} = V_{dd} I_{dd} = \frac{V_{dd} \cdot I_{\max}}{\pi} \cdot \frac{(\sin\theta_0 - \theta_0\cos\theta_0)}{1-\cos\theta_0} \tag{5.52}$$

$$P^C_{RF,f_0,\max} = \frac{v_{d,1} i_{d,1}}{2} = \frac{I_{\max} \cdot (V_{dd} - V_{knee})}{2\pi} \cdot \frac{(\theta_0 - \sin\theta_0\cos\theta_0)}{1-\cos\theta_0} \tag{5.53}$$

$$\eta^C_{D,\max} = \frac{P^C_{RF,f_0,\max}}{P^C_{dc,\max}} = \frac{(V_{dd} - V_{knee})}{2V_{dd}} \cdot \frac{(\theta_0 - \sin\theta_0\cos\theta_0)}{(\sin\theta_0 - \theta_0\cos\theta_0)} \tag{5.54}$$

if $V_{knee} = 0$, then $\eta^C_{D,\max} = \frac{1}{2}\frac{(\theta_0 - \sin\theta_0\cos\theta_0)}{(\sin\theta_0 - \theta_0\cos\theta_0)}$

The dissipated power in the transistor for class C operation mode can be calculated as:

$$P_{diss}^{B} = \frac{1}{2\pi}\int_{0}^{2\pi} V_{ds}(\theta)\cdot I_{ds}(\theta)d\theta = P_{dc} - P_{RF} = \frac{V_{\max}\cdot I_{\max}}{4\pi}\cdot\frac{2\sin\theta_0 + \sin\theta_0\cos\theta_0 - 2\theta_0\cos\theta_0 - \theta_0}{1-\cos\theta_0} \quad (5.55)$$

5.3.2 Switching-Mode PAs (Classes D, E, F)

In contrast to class A, AB, B and C PAs, where operation in the triode region should be avoided, class D, E, and F PAs rely on operation in the triode region for optimal efficiency and output power. In these amplifiers, which are also called switching-mode amplifiers, the output device is driven by a large square wave signal. For signals that have mainly phase and frequency modulation, such as quadrature phase-shift keying (QPSK) or Gaussian minimum-shift keying (GMSK), the envelope is constant. This means that a nonlinear high-efficiency PA can be used to amplify such signals.

The current and voltage waveforms at the drain level of the transistor can be shaped to reduce the power dissipation and enhance the efficiency of PAs. This reduction can be achieved by avoiding simultaneous presence of high current and high voltage at the drain level of the transistor, since the power dissipation is equal to the product of both current and voltage.

In the ideal case, no power dissipation occurs if the transistor behaves as a lossless switch, which makes the transistor operate in two states. In the first state, when the transistor is on, the voltage across it is zero and the current is high, i.e., the transistor behaves as a closed switch during this part of a cycle. In the second state, when the transistor is off, its current is zero and its voltage is high, i.e., the transistor behaves as an open switch during this part of a cycle.

The gate and drain voltages and drain current waveforms for an ideal switch-like transistor are shown in Figure 5.14. One way to achieve the switching behavior is proper termination of the device at the fundamental and harmonic frequencies, in order to shape the current and voltage waveforms to the switching-mode operation [9].

The drain voltage and current corresponding to the waveforms in Figure 5.14 can be described as:

$$I_{ds}(\theta) = \frac{I_{\max}}{2} + \sum_{n=0}^{\infty}\frac{2I_{\max}}{(2n+1)\pi}\sin(2n+1)\theta \quad (5.56)$$

$$V_{ds}(\theta) = \frac{V_{\max}}{2} + \sum_{n=0}^{\infty}\frac{2V_{\max}}{(2n+1)\pi}\sin(2n+1)\theta \quad (5.57)$$

where $\frac{V_{\max}}{2} = V_{dd}$.

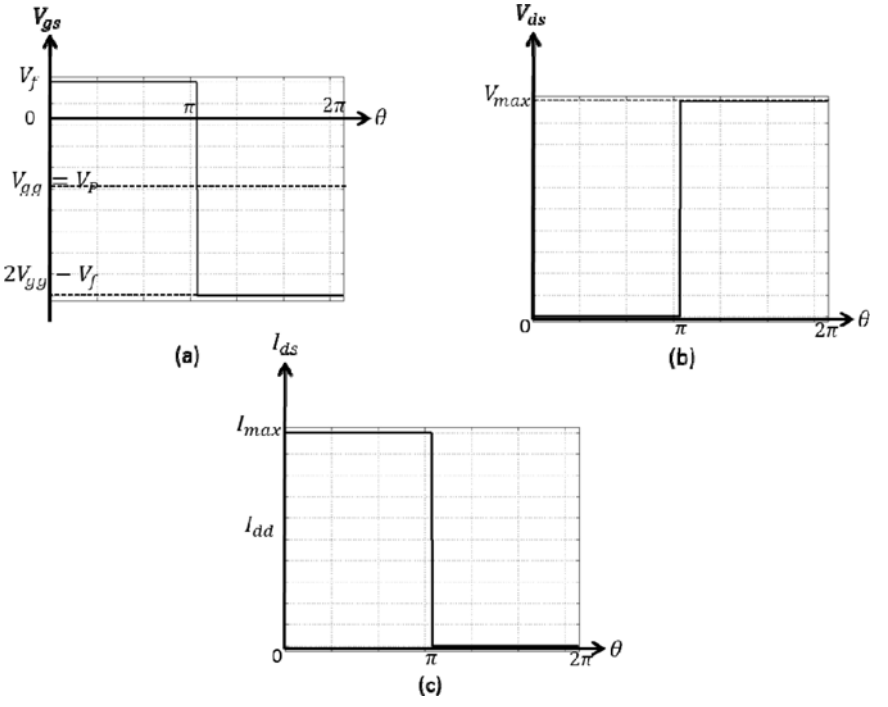

Fig. 5.14 Ideal switch-like transistor: (a) gate voltage, (b) drain voltage and (c) drain current waveforms

The dissipated power in the transistor for ideal switching-mode operation is calculated as:

$$P_{diss}^{B} = \frac{1}{2\pi}\int_{0}^{2\pi} V_{ds}(\theta)\cdot I_{ds}(\theta)d\theta = P_{dc} - P_{RF} = 0 \tag{5.58}$$

This is because the multiplication of the drain voltage and drain current is zero at each angular phase. Hence, all the applied DC power is transfer to RF output power. The DC power is calculated as:

$$P_{dc} = \frac{V_{\max} I_{\max}}{4} \tag{5.59}$$

However, practically, the RF output power of the transistor consists of the fundamental frequency power, desired output power, and unwanted harmonic frequencies power. The output RF power for different harmonic frequencies can be obtained as:

$$P_{RF} = \sum_{n=1}^{\infty}\frac{1}{2}\left|V_{ds,nf_0} I_{ds,nf_0}\right| = \sum_{n=0}^{\infty}\frac{2V_{\max} I_{\max}}{(2n+1)^2\pi^2} \tag{5.60}$$

where V_{ds,nf_0} and I_{ds,nf_0} are the drain voltage and current amplitudes of the harmonic signals. Simply, the following equation can be obtained:

$$P_{RF} = \frac{V_{\max} I_{\max}}{4} \tag{5.60}$$

As a result, the load harmonic impedance network should be organized such that unwanted harmonic power is rejected, while the drain waveforms are controlled to keep the dissipated power in the transistor at zero. Consequently, all the DC power is transferred to the fundamental frequency output power, and the output power efficiency theoretically becomes 100%.

The switch-like behavior of the PA is administered by the gate biasing and the RF input signal. The fundamental and harmonic load impedances supplied for the drain of the transistor shape the drain waveforms, which define the different classes of switching-mode PAs. Switching-mode PAs can be categorized into two groups as:

- Single branch configurations, such as classes E and F, and inverse class F PAs; and,
- Balanced or push-pull configurations, such as voltage mode class D and current mode class D PAs.

5.3.2.1 Class D PA

The increased efficiency of class D PAs is a result of exploiting transistors as switches. In most cases, a class D PA implements a pair of active devices operating in a push-pull mode and a tuned output circuit. Switching-mode PAs with an output filter tuned to the fundamental frequency ideally transform all the DC power to fundamental frequency power, which can be delivered to the load without power losses at the harmonics. This results in a power output of $(8/\pi^2)V_{dd}^2/R$ for the transformer-coupled configuration. The output circuit is tuned to the switching frequency and ideally removes its harmonic components, which results in a completely sinusoidal signal that can be delivered to the load. Current is drained only through the transistor that is on, resulting in a 100% efficiency for an ideal class D PA [6], [10].

Class D amplifiers can be categorized into two groups as:

- Class D voltage-switching PAs (also designated as voltage mode class D, VMCD); and,
- Class D current-switching PAs (also designated as current mode class D, CMCD).

Figure 5.15 shows a circuit schematic of a voltage-switching push-pull class D PA, where C_b is the bypass capacitor, C_o is the blocking capacitor, and R is the load resistance. By using the input transformer, both transistors can be driven with currents that are 180° out of phase. This can be done by reversing one secondary winding on the transformer. The RF connection of the transistor output is in parallel configuration as a result of the grounding effect of a bypass capacitor, C_b. Hence, the equivalent load resistance is equal to $2R$ for each device.

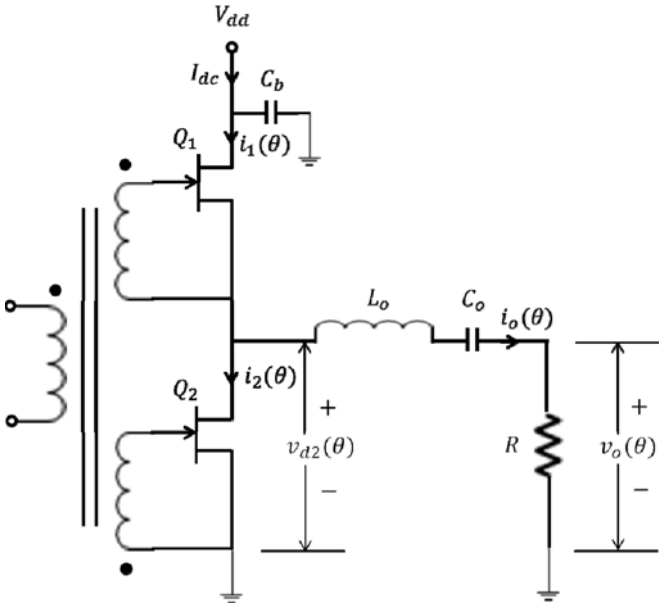

Fig. 5.15 Voltage-switching push-pull class D PA

The voltage and current waveforms are shown in Figure 5.16 for a voltage-switching push-pull class D PA. The square voltage waveform enables good utilization of the device's breakdown capability. The maximum amplitude of the half sinusoidal current through each transistor is $I_{\max} = \pi \cdot I_{dc}$.

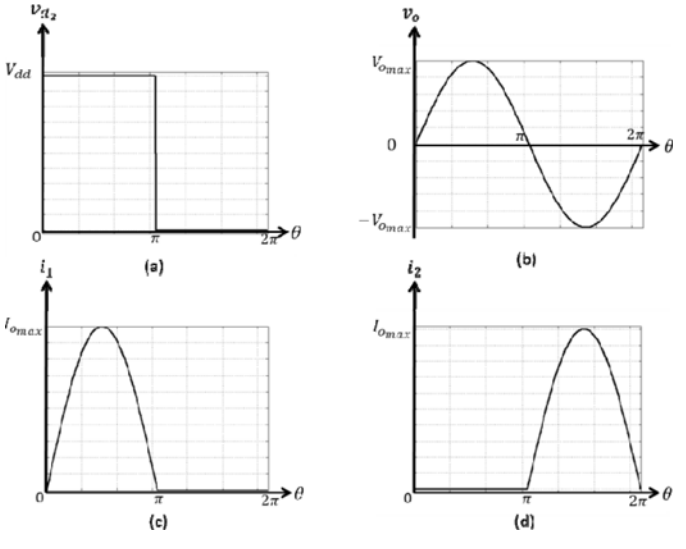

Fig. 5.16 Voltage-switching class D voltage and current waveforms

The Fourier series of current waveforms, $i_1(\omega t)$ and $i_2(\omega t)$, can be obtained as:

$$i_1(\omega t) = I_{dc}\left(1+\frac{\pi}{2}\sin(\omega t)-2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2-1}\right) \tag{5.61}$$

$$i_2(\omega t) = I_{dc}\left(1-\frac{\pi}{2}\sin(\omega t)-2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2-1}\right) \tag{5.62}$$

The Fourier series of the voltage at a connecting node can be obtained as:

$$v_{tot}(\omega t) = V_{dd}\left(\frac{4}{\pi}\right)\left(\sum_{n=1,3,5,\ldots}^{\infty}\frac{\sin n\omega t}{n}\right) \tag{5.63}$$

The current passing through the load is given by:

$$i_R(\omega t) = i_1(\omega t) - i_2(\omega t) = \pi \cdot I_{dc} \cdot \sin(\omega t) \tag{5.64}$$

Voltage-switching class D PAs are widely used in audio applications. However, they have been rarely used at frequencies in the range of 1 GHz or higher. This is because of the losses associated with the transition time, parasitic reactance and turn-on resistance of the transistors. At high frequencies, the losses in the output parasitic reactance, mainly the shunt capacitance, C_{ds}, dominates the loss due to the turn-on resistance. The energy, E, dissipated within C_{ds} per switching cycle can be obtained by:

$$E = \frac{1}{2}C_{ds} \cdot V_{dd}^2 \tag{5.65}$$

where V_{dd} is the voltage across the transistor at switch closure [11].

To overcome this limitation, current-switching class D PAs have been proposed. In this configuration, a shunt LCR (inductor, resistor, capacitor) output filter is implemented. In fact, the current waveform becomes a voltage waveform and vice versa. A series connection transforms to a parallel connection and vice versa. Furthermore, a voltage-source supply to transformer center-tap converts to a RF-choke feed (current source) to transformer center-tap.

The use of the current-switching class D PA cancels out the losses in the shunt capacitance by achieving a zero voltage switching. Therefore, current-switching class D amplifiers are more appropriate for high-frequency applications resulting in higher power efficiency.

Figure 5.17 shows a current-switching class D PA, in which the two devices are connected in parallel. An input transformer is required to drive the devices out-of-phase, and also an output transformer is required to extract the differential load

voltage. As demonstrated in Figure 5.18, the currents through each transistor have the shape of square waveforms. In order to have proper switching, large input power is required. Voltages v_1 and v_2 have the shape of half-sinusoidal waveforms. The maximum amplitude of the half–sinusoidal voltage waveforms is $V_{\max} = \pi \cdot V_{dd}$, and high breakdown voltage is required.

The Fourier series of voltage waveforms, $v_1(\omega t)$ and $v_2(\omega t)$, can be calculated as:

$$v_1(\omega t) = V_{dd}\left(1 + \frac{\pi}{2}\sin(\omega t) - 2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2 - 1}\right) \tag{5.66}$$

$$v_2(\omega t) = V_{dd}\left(1 - \frac{\pi}{2}\sin(\omega t) - 2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2 - 1}\right) \tag{5.67}$$

The Fourier series of the RLC tank current waveform is given by:

$$I_{tot}(\omega t) = I_{dc}\left(\frac{4}{\pi}\right)\left(\sum_{n=1,3,5,\ldots}^{\infty}\frac{\sin n\omega t}{n}\right) \tag{5.68}$$

The voltage across the load is obtained as:

$$v_{load}(\omega t) = v_1(\omega t) - v_2(\omega t) = \pi \cdot V_{dd} \cdot \sin(\omega t) \tag{5.69}$$

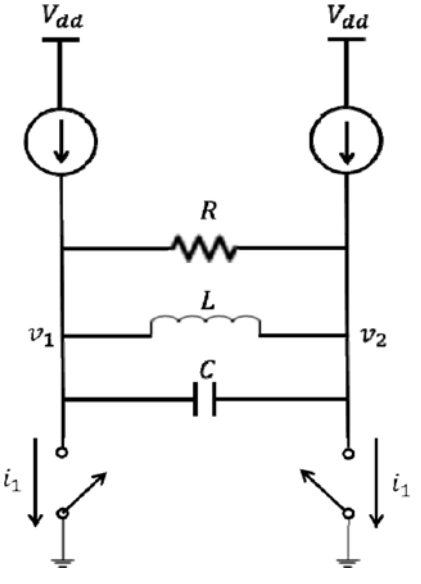

Fig. 5.17 Current-switching class D PA

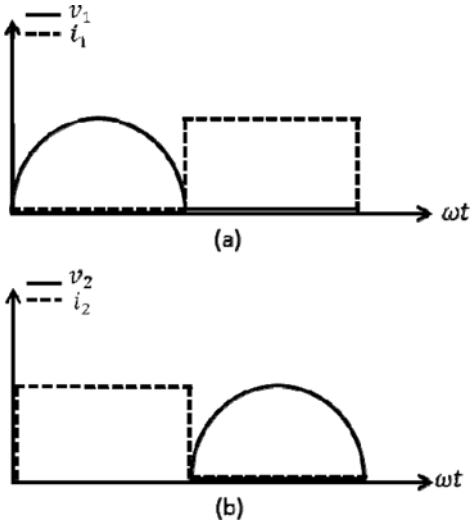

Fig. 5.18 Current-switching class D voltage and current waveforms

5.3.2.2 Class E PA

In a class E PA, a single transistor operates as a switch. The drain voltage waveform results from the sum of the DC and RF currents charging the drain current capacitance. In an ideal class E PA, the drain voltage drops to zero and has a zero slope when the transistor turns on (zero voltage switching). This eliminates the losses associated with charging the drain capacitance in class D PAs and considerably decreases the switching losses [12].

In order to optimize the efficiency, the transient response of the load network has to be controlled even at the time the switching speed is a considerable fraction of the RF cycle. Indeed, class E PAs can be designed to operate at frequencies higher than those of class D PAs. Class E PAs can be designed at frequency bands up to 3.5 GHz, while attaining high power efficiency.

In a switching-mode class E PA, it is possible to have simultaneous high voltage and high current during the switching between on and off. Therefore, a switching-mode PA with resistive load has power dissipation during switching transitions. Class E can have good efficiency only if switching times are much smaller than the waveform period. Figure 5.19 shows current and voltage waveforms in a square-wave switching-mode class E amplifier with resistive load.

Switching power dissipation at every instant of time can be obtained as the multiplication of the voltage and current in the active part of the transistor. The current and voltage must rise and fall, but not at the same time. In other words, when the voltage rises, the current must fall and vice versa. Furthermore, time-displace voltage and current transitions never have high current and high voltage at the same time. In a class E PA, switching has to be performed as fast as possible with a reasonable input drive. A class E PA has been utilized for high-efficiency amplification at frequencies up to K-band [6].

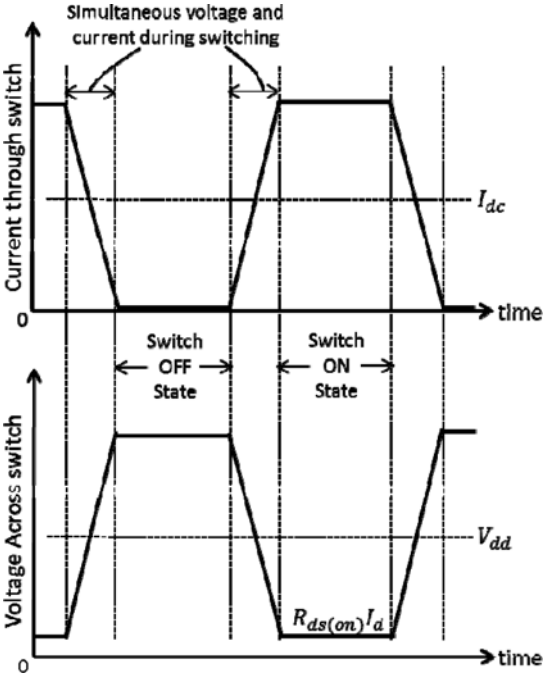

Fig. 5.19 Current and voltage waveforms in a square-wave switching-mode class E PA

5.3.2.3 Class F and Inverse Class F PAs

A class F PA is characterized by a load network that has resonances at one or more harmonic frequencies as well as at the carrier frequency. Class F amplifiers are capable of high efficiency (88.4% for traditionally defined class F or 100% if infinite harmonic tuning is used), which results from a low DC voltage current product. In other words, the drain voltage and current are shaped to minimize their overlap region.

Class F amplifier design is challenging mainly due to the complex design of the output-matching network [6]. At microwave frequencies, it becomes difficult to fabricate class F amplifiers, as capacitors and inductors function poorly; and, planar structures becomes hard to realize, especially when tuning multiple harmonics.

If the even harmonics are terminated to short circuits and the odd harmonics are terminated to open circuits, the current and voltage waveforms have the shapes of half-sinusoidal and square waveforms, respectively. The Fourier series of half-sinusoidal current and square voltage waveforms, both at the same angular velocity, ω, are given by:

$$i(\omega t) = I_{dc}\left(1 - \frac{\pi}{2}\sin(\omega t) - 2\sum_{n=2,4,6,\ldots}^{\infty}\frac{\cos n\omega t}{n^2 - 1}\right) \tag{5.70}$$

$$v(\omega t) = V_{dc}\left(1+\frac{4}{\pi}\sin(\omega t)+\frac{4}{\pi}\sum_{n=3,5,7,\ldots}^{\infty}\frac{\sin n\omega t}{n}\right) \tag{5.71}$$

where I_{dc} and V_{dc} are the DC current and voltage components, respectively [13].

This class of operation is called class F or voltage mode class F. The voltage and current waveforms of the theoretical class F PA is demonstrated in Figure 5.20 (a). The power dissipation for this class of operation, which is equal to the product of the current and voltage, is zero. Therefore, a class F PA is able to achieve 100% efficiency in theory.

Inverse class F or current mode class F PA operation is obtained by terminating the odd harmonics to short circuits and the even harmonics to open circuits. In this case, the current and voltage waveforms have the shapes of square and half-sinusoidal waveforms, respectively. The Fourier series of current and voltage waveforms are in the form of:

$$i(\omega t) = I_{dc}\left(1+\frac{4}{\pi}\sin(\omega t)+\frac{4}{\pi}\sum_{n=3,5,7,\ldots}\frac{\sin n\omega t}{n}\right) \tag{5.72}$$

$$v(\omega t) = V_{dc}\left(1-\frac{\pi}{2}\sin(\omega t)-2\sum_{n=2,4,6,\ldots}\frac{\cos n\omega t}{n^2-1}\right) \tag{5.73}$$

The voltage and current waveforms of the theoretical inverse class F are shown in Figure 5.20 (b). As in class F operation, the power dissipation of inverse class F operation is zero; therefore, this class is also capable of achieving 100% efficiency in theory [13].

However, due to some limitations, the theoretical 100% efficiency of class F and inverse class F PAs is not possible in practice. In general, an infinite number of serial odd-harmonic tank resonators can provide a half-sinusoidal current waveform and a square voltage waveform. Figure 5.21 (a) shows such an output-matching network with a multiple-resonator output filter that tunes the harmonics to the class F operational conditions. All the even harmonics are tuned with one parallel tank resonating at the fundamental frequency. Likewise, the harmonic matching for the inverse class F operation needs an infinite number of serial even-harmonic tank resonators and one parallel tank resonating at the fundamental frequency. The harmonic matching network for the inverse class F is shown in Figure 5.21 (b). Indeed, such a harmonic matching circuit is not feasible in practice.

Berini et al. [14] and Raab [15] showed that the significant effect on the waveform shaping and, in fact power efficiency, is only related to the first few harmonics. Tuning more harmonics slightly enhances the efficiency of the PA at the cost of increasing design complexity and output matching losses. In addition to the matching networks' complexity and losses, the intrinsic and extrinsic effects of the transistor impact the maximum efficiency that can be achieved.

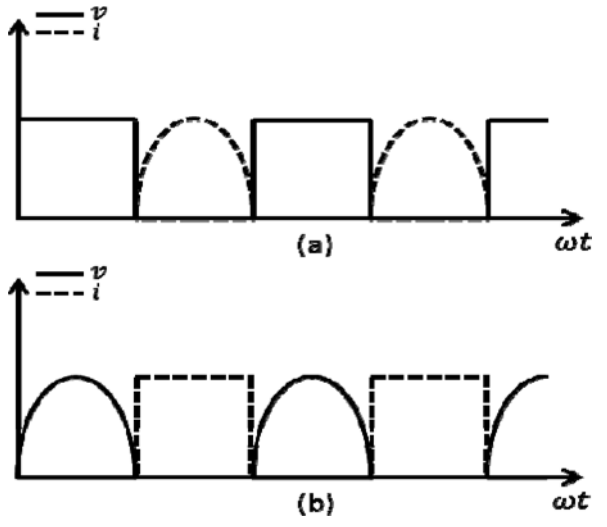

Fig. 5.20 Ideal waveforms in switching mode PAs: a) class F and b) inverse class F

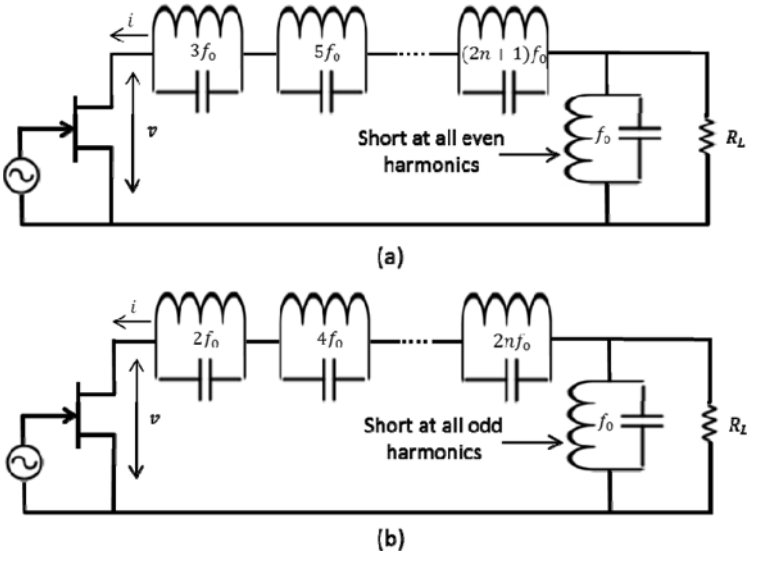

Fig. 5.21 Basic circuits of PAs: a) class F and b) inverse class F

5.3.3 Comparison of the Classes of Operation in RF PAs

Table 5.1 compares the performance of the different classes of operations described in the preceding sections, in terms of output power, gain, efficiency, and linearity. Table 5.1 shows that the efficiency of the linear amplifiers decreases from class A to class C. However, the high linearity of the class A PA trends to the high nonlinearity of the class C PA when moving from class A to class C. In the case of switching-mode PAs, class D and E PAs have very high efficiencies, but they are strongly nonlinear. The class F PA also has very high efficiency and is highly nonlinear.

Table 5.1 Performance Comparison for Different Class of Operations of PAs

	Class A	Class AB	Class B	Class C	Class D	Class E	Class F
Output power	++++	+++	++	+	++	++	+++
Gain	++++	+++	++	+	++	++	+++
Efficiency	+	++	+++	++++	++++	++++	+++
Linearity	++++	+++	++	+	+	+	++

5.4 Linearization of RF Power Amplifiers

The nonlinear behavior of the RF front-end, especially RF transmitters, can significantly degrade the overall performance of wireless systems. The power efficiency of an RF amplifier is optimal when it is operated near saturation [16]. An amplifier operating in this nonlinear range generates IM distortion that interferes with neighboring channels. Therefore, there should be compensation for the nonlinearities and distortions of the RF transmitter.

An efficient PA design reduces the cost of power consumption and increases the battery life of wireless mobile transmitters. However, the increase in efficiency is usually accompanied by linearity deterioration, which requires a trade-off between efficiency and linearity.

The output power, gain and efficiency variations for a typical PA as functions of the input power are shown in Figure 5.22. As can be observed from this figure, there are two major operational regions for the amplifier. In the first region, the amplifier has a linear gain, but the efficiency is low. In the second region, the amplifier has high efficiency, but the gain is nonlinear. Linearity deterioration is caused by the distortions introduced to the signal by the compression of the PA in the saturation region. As a result, the linearity/nonlinearity of the PA depends on the input signal power [17].

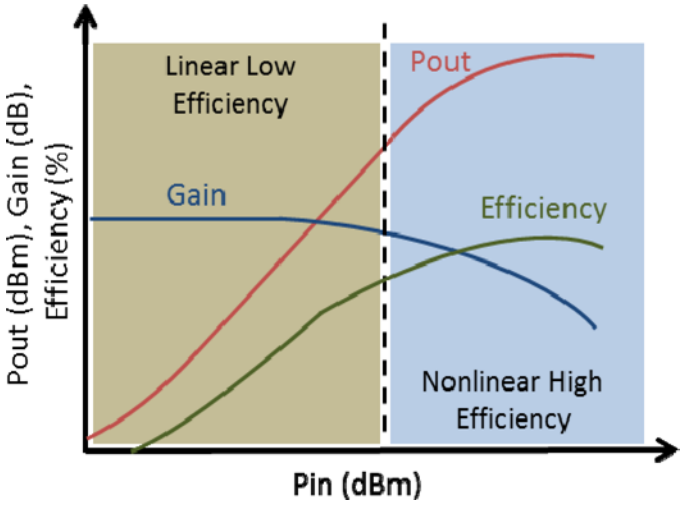

Fig. 5.22 Output power, gain and efficiency of a PA versus input power

Linearization is a systematic approach to reduce an amplifier's distortion and is inevitable for enhancing the linearity of an amplifier to the high input power drive levels and achieving linearity requirements when operating the device over its entire power range [18]. Linearization allows a PA to generate more power and operate with higher efficiency for a given level of distortion. There are different methods for linearizing an RF amplifier. The three major linearization techniques are feedback, feedforward, and predistortion [19].

5.4.1 Feedback Linearization

The feedback linearization technique is a closed-loop system that can provide high levels of linearization. It is based on the concept of the feedback loop, which is widely used in control theory. As shown in Figure 5.23, the amplifier's output signal is fed back and subtracted from the amplifier's input signal to force the output signal to be a linear replica of the input signal. The use of feedback linearization has been widely investigated, but has received little use at RF frequencies. The most important reason is probably related to issues with amplifier stability and the difficulty in production of networks with non-ideal components that function over wide frequency ranges [19].

Indirect feedback techniques are reported to be more widely used. The correction in two alternative indirect feedback strategies, Cartesian feedback and polar feedback, are processed in a baseband domain. Two loops are needed for the amplitude and phase. If only one loop is used for the amplitude at lower frequencies, the technique is also called an automatic gain control (AGC).

There are some limitations with feedback linearization techniques that make them unsuitable for wireless communication applications. The gain of the linearized amplifier is reduced to the gain of feedback loop. Furthermore, the delays associated with the feedback loop must be small enough to ensure stability, which limits the use of feedback linearization techniques to narrowband signals [20].

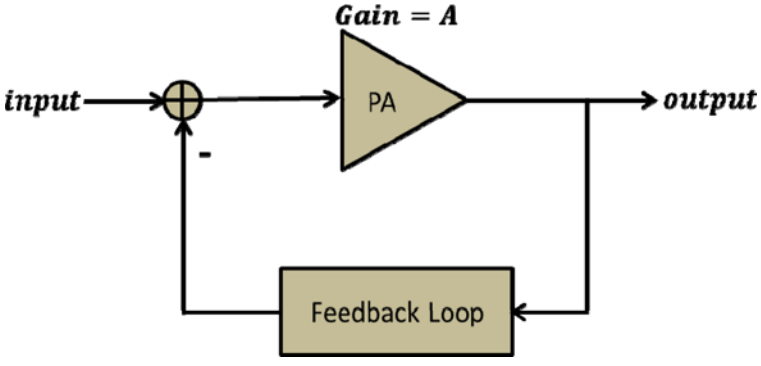

Fig. 5.23 Feedback linearization block diagram

The frequency bandwidth of feedback linearization is limited by the inverse of the propagation delay of the feedback loop. For third-order distortion correction, the bandwidth of the loop must be three times the bandwidth of the baseband envelope. Moreover, at small signals, the system becomes inefficient, and the modulator/demodulator has a limited dynamic range. Operational amplifiers and demodulators introduce noise in the system. Stability, noise and orthogonality of the modulator/demodulator can be improved if the loop is digitized with a digital signal processor (DSP). However, current DSPs cannot reach the speed of the equivalent analog systems.

Different approaches have been proposed to address the limitations, including envelope feedback, polar feedback, Cartesian feedback, adaptive double envelope feedback, and digital Cartesian feedback. In envelope feedback linearization, input and output samples are fed into two envelope detectors. As illustrated in Figure 5.24, the error between the outputs of the envelope detectors is injected to the main signal path via a vector modulator. The use of this error signal to control the amplifier's output only allows the compensation for AM/AM nonlinearities of the PA, not its AM/PM nonlinearities.

Hence, a second feedback loop is added in polar feedback linearization, which is shown in Figure 5.25. In polar feedback linearization technique, the addition of phase lock loop (PLL) to the envelope feedback linearizer allows the compensation for both AM/AM and AM/PM nonlinearities of the PA.

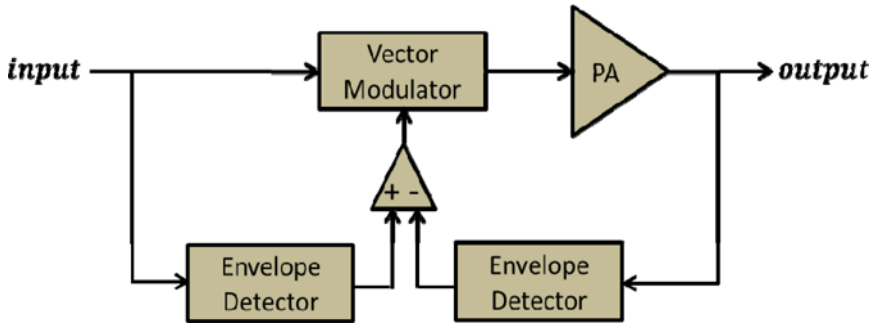

Fig. 5.24 The envelope feedback linearization technique

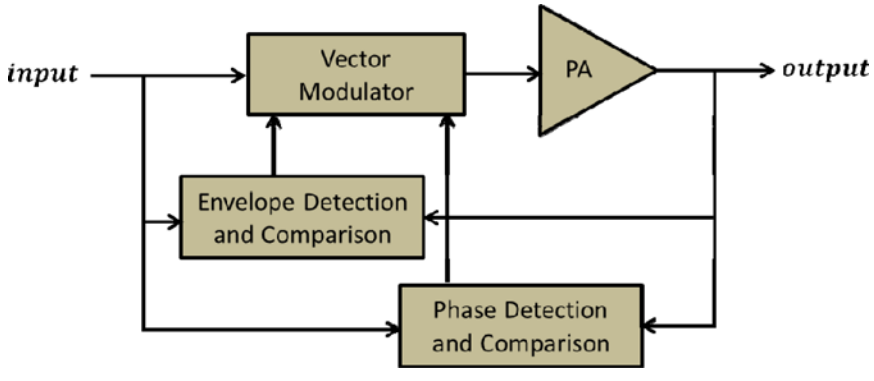

Fig. 5.25 The polar feedback linearization technique

An alternative approach is Cartesian feedback, which separates the signal into in-phase and quadrature components. In this technique, the need for phase-shift components is eliminated; and, correction of gain and phase is still possible by tuning the amplitudes of two orthogonal components. The baseband in-phase and quadrature components are compared to control the attenuators in the vector demodulator [19].

5.4.2 Feedforward Linearization

Feedforward linearization is a powerful technique that has the same fundamental error-correcting operation as feedback, but removes the inherent bandwidth and stability problems of a conventional feedback loop at RF frequencies. The feedforward technique is theoretically correct for all nonlinearity orders.

In feedforward linearization, a correction signal is injected at the output of the PA to cancel its nonlinearity distortions. Figure 5.26 demonstrates the block

diagram of the feedforward linearization technique. As can be observed from this figure, there are two loops in the feedforward system, which are the signal cancellation loop and the distortion cancellation loop [6]. In the signal cancellation loop, a portion of the PA's output signal is compared to a time-, amplitude- and phase-aligned replica of the input signal. The signal at the output of the comparator represents the additive distortion products at the output of the PA. These products, after being amplified to their original magnitudes, are then subtracted from the time-aligned version of the PA output, which theoretically leads to perfect cancellation of the distortions generated by the PA [2].

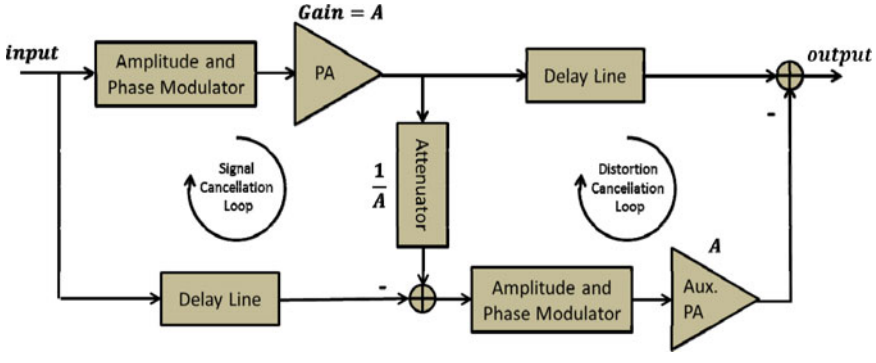

Fig. 5.26 Feedforward linearization block diagram

The feedforward linearization technique is naturally sensitive to changes in operating conditions. However, the development of adaptation methods to compensate for such changes has increased the interest in the technique. Due to the phase shifts introduced in the carrier by the amplifiers, the feedforward linearizers are very sensitive to the alignment in both loops [16]. Indeed, delay lines must be precisely tuned in order to achieve desirable performance. Furthermore, feedforward linearization is intrinsically non-adaptive and requires adaptive control techniques in both loops to retain the proper alignment. The gains of the attenuator and auxiliary PA must be very well matched to the gain of the PA. Furthermore, the delay line and subtractor in the output path of the PA have to be low loss to achieve an effective linearization.

Feedforward linearizers are commonly used in wireless communication base stations. In these linearizers, very high linearity can be achieved (50 dBc or higher), and wide bandwidth can be supported (40-60 MHz or more). Power efficiency is the major drawback of this technique. Although the main PA is operating at high efficiency, the overall efficiency of the linearized amplifier is considerably decreased by the error amplifier, which needs to be perfectly linear. The overall efficiencies of third-generation (3G) multicarrier PAs linearized by the feedforward technique are in the range of 10%-15%.

5.4.3 Digital Predistortion Linearization

The predistortion can be performed in the analog or digital domain and on baseband, intermediate frequency (IF) or RF signals. Digital baseband predistortion, which is an open-loop technique, is currently the preferred linearization technique and is widely used for applications up to 20 MHz bandwidth. In a digital predistortion (DPD) technique, a complementary nonlinearity upstream of the PA is applied, so that the cascade of the digital predistorter and the PA behaves as a linear amplification system. Figure 5.27 presents a simplified block diagram of the DPD linearization technique [4]. DPD achieves high linearity performance with high power efficiency. The efficiencies of 3G PAs linearized by digital predistorters are reported in the range of 30-45%.

Behavioral modeling of RF PAs is an essential task in the design of high-performance wireless transmitters that combine high power efficiency and spectrum compliant linearity performance. This is even more important in modern communication systems that employ envelope-varying signals with high PAPRs. Indeed, such signals set rough linearity requirements on the PA and usually lead to the forfeit of the system's power efficiency, in order to meet the linearity requirements [1].

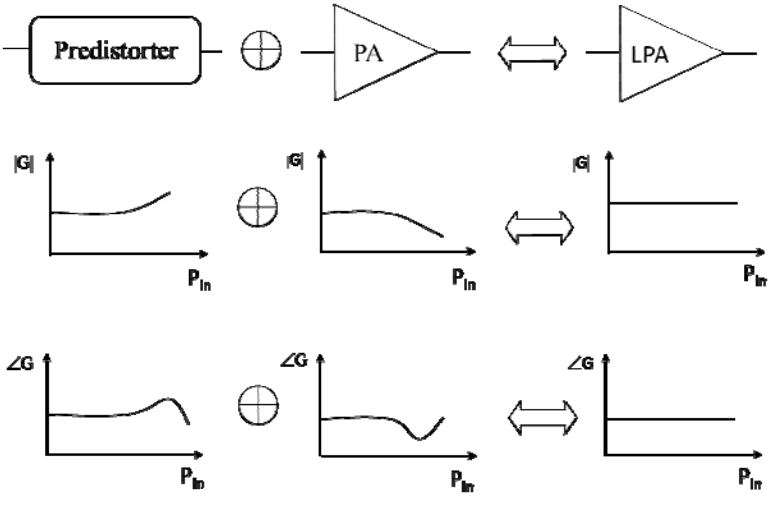

Fig. 5.27 Predistortion linearization technique

The most important advantage of behavioral modeling is that it does not require extensive knowledge of the RF circuit physics and functionality. As demonstrated in Figure 5.28, behavioral modeling simplifies the modeling of the RF circuit to the identification of a mathematical formulation that relates the input and output of the device under test (DUT) that can be considered as a black box. Consequently,

behavioral modeling appears as a time and resource efficient process for transmitter performance evaluation and digital predistorter design [4].

Due to current broadband and highly varying signals, such as code-division multiple-access (CDMA), wideband CDMA (WCDMA), andorthogonal frequency-division multiplexing (OFDM), PAs or wireless transmitters have to be considered as dynamic nonlinear systems. Therefore, an appropriate architecture has to be chosen to extract and identify an accurate and robust forward or reverse model for such systems [1].

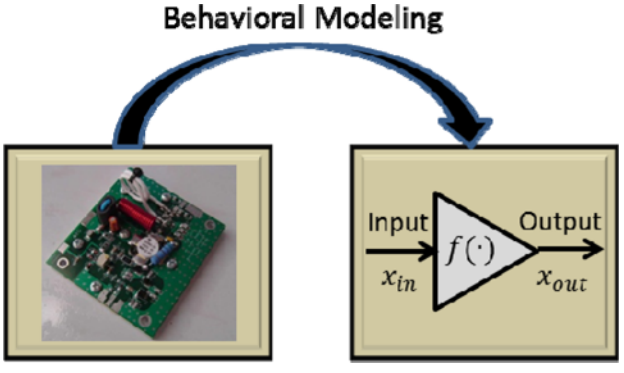

Fig. 5.28 Black-box based behavioral modeling

The block diagram of a DPD-based linear amplification system is shown in Figure 5.29. The signal at the input of the digital predistorer ($x_{in\_DPD}(n)$) can be derived from the signal at the output of the PA block according to:

$$x_{in\_DPD}(n) = \frac{x_{out}(n)}{G} \tag{5.74}$$

The behavioral modeling of the DUT identifies function f_{DUT} that satisfies $f_{DUT}(x_{in}(n)) = x_{out}(n)$. The synthesis of the DPD function is equivalent to the estimation of function f_{DPD} , such that $f_{DPD}(x_{out}(n)/G) = x_{in}(n)$.

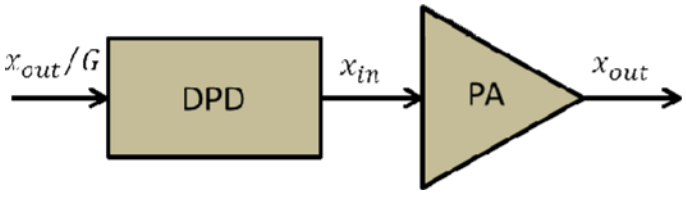

Fig. 5.29 Block diagram of digital predistortion (DPD) based linear PA

Predistorters potentially can handle much wider modulation bandwidths, including multicarrier signals. Unlike closed-loop systems, they do not have stability problems. Predistorters can be integrated in monolithic microwave integrated circuits (MMICs). They provide good linearity improvement in traveling tube wave amplifiers (TWTAs), but less significant improvement for metal semiconductor field-effect transistor (MESFET) classes A and AB. Predistortion linearization techniques have a smaller dynamic range than feedforward linearization techniques. In addition, predistorters require temperature compensation or an adaptive control mechanism.

Numerous formulations have been proposed for behavioral modeling and DPD of RF PAs and transmitters. Some commonly used formulations are memoryless look-up table, nested look-up table, Volterra, memory polynomial, forward, reverse and parallel twin nonlinear two-box, envelope memory polynomial, Wiener and augmented Wiener, and Hammerstein and augmented Hammerstein models. The look-up table and memory polynomial models are described in the following sections. More information about the other models can be found in [4].

5.4.3.1 Look-Up-Table Model

The static look-up table (LUT) model is the basic behavioral model for memoryless AM/AM and AM/PM nonlinearities. The complex gain of the DUT is stored in two look-up tables. The output waveform can be shown as:

$$x_{out}(n) = G\left(\left|x_{in}(n)\right|\right) \cdot x_{in}(n) \tag{5.75}$$

where $G\left(\left|x_{in}(n)\right|\right)$ is the instantaneous complex gain of the DUT.

The AM/AM and AM/PM characteristics of the DUT can be obtained from the raw measured data using averaging or polynomial fitting techniques.

5.4.3.2 Memory Polynomial Model

The memory polynomial model is widely used for behavioral modeling and DPD of PAs and transmitters exhibiting memory effects. The output waveform of memory polynomial model is given by:

$$x_{out}(n) = \sum_{j=0}^{M} \sum_{i=1}^{N} a_{ji} \cdot x_{in}(n-j) \cdot \left|x_{in}(n-j)\right|^{i-1} \tag{5.76}$$

where N and M are the nonlinearity order and the memory depth of the DUT, respectively; and, a_{ji} are the model coefficients.

Figure 5.30 demonstrates typical spectrum results for a linearized Doherty amplifier using a memory polynomial based digital predistorter. This figure clearly presents the perfect cancellation of the nonlinearities introduced by the DUT.

There are other variations of the memory polynomial model in the literature, including the orthogonal memory polynomial model and the memory polynomial model with cross-terms, which is also called as the generalized memory polynomial model [4].

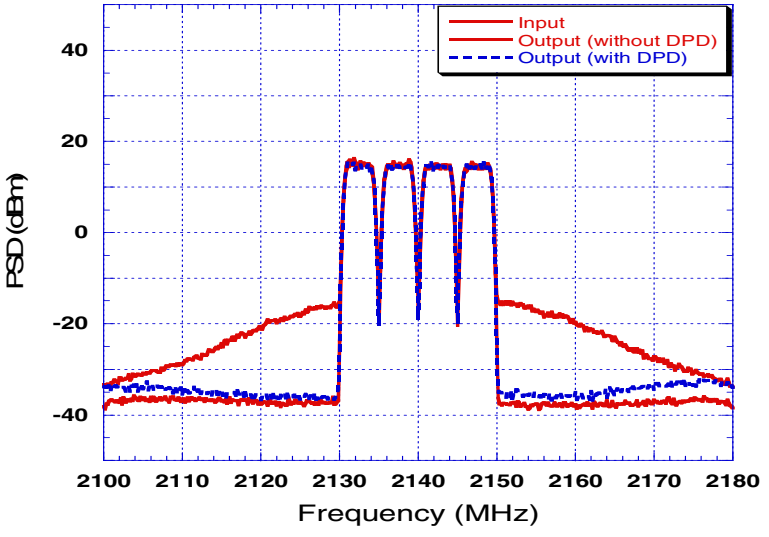

Fig. 5.30 Performance of a DPD linearized 30-Watt GaN Doherty PA with four-carrier WCDMA signal (PAPR = 11.25 dB) [4]

5.4.4 Analog Predistortion Linearization

RF linearizers based on predistortion have been used for several years in satellite transponder PAs, either as solid-state power amplifiers (SSPAs) or TTWTAs, to correct the nonlinear characteristics. The use of predistorters together with the PAs on board a satellite allows for the operation of its transponders at reduced distortion levels and higher power efficiency, while carrying multicarrier traffic [21]. Analog predistorters (PDs) are still in use for high power applications in the upper GHz frequency bandwidths, where medium linearization performance is targeted. PDs have less complexity, low implementation cost, wide bandwidth and the capability to be added to existing PAs as a separate stand-alone component. Applications include mobile and handset PAs, which usually requires a simple PD network, typically a diode gain amplifier [7].

A linearizer based on a series diode with a parallel capacitor is a very simple structure that utilizes the nonlinearity of the series resistance of the diode, which produces a characteristic positive gain and negative phase with increasing input power. This is a simple linearizer, but requires an additional isolation mechanism between the linearization circuit and the PA. Furthermore, it has very limited control on the achieved characteristics and, hence, has very limited practical applications [22].

The most basic linearization circuit with diodes is shown in Figure 5.31 (a) in a head-tail configuration, which consists of semiconductor diodes as a nonlinear resistor. This setup suffers from linear phase distortions in the AM/PM characteristic

of a PA. The linearizer can be tuned to resonate at a center frequency to mitigate this effect, but the linearizer is a nonlinear circuit. Therefore, the resonance cannot be achieved for the full range of input signals. Furthermore, at microwave frequencies, the problem is even more severe, as the capacitive effects of the junction capacitance of the Schottky diodes limit the application of this linearizer to a very moderate range of input signals. This has been addressed in [22].

To overcome this problem, the resonance can be attained by utilizing a suitable inductance, L_s in shunt, as illustrated in Figure 5.31 (b). By utilizing this shunt inductance, the problem of phase distortion can be reduced for any input power and any bias voltage, V_{sch}, of Schuttky diodes D_1 and D_2. The actual bias voltage V_{PK} across the Schottky diodes D_1 and D_2 is regulated by the bias feed resistance, R_b, and the input RF power. The PIN diode, D_3, delivers a variable resistance in shunt to the biased Schottky diodes D_1 and D_2. D_3 provides a dynamic control to achieve the required gain and phase characteristics and also provides an improved dynamic range in the gain and phase responses, V_{pin}. The actual bias of the PIN diode is regulated by the bias feed resistance, R_{b1}, and the input RF power.

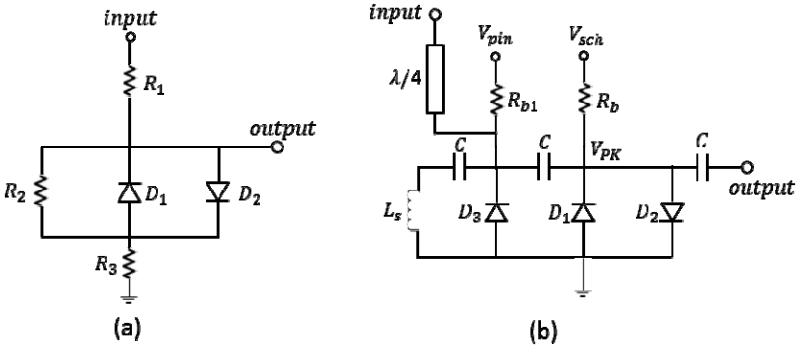

Fig. 5.31 (a) The basic linearization circuit with diodes, (b) an analog linearization circuit

PAs may have single or dual inflection points in their distortion characteristics based on device technology or bias conditions, as shown in Figure 5.32. A diode-based RF PD linearizer that can be used to linearize PAs with dual inflection points in their distortion characteristics is proposed in [23]. The block diagram of the suggested PD is shown in Figure 5.33. Two linearizer branches connected to through and coupled ports of a branch line hybrid coupler. Each linearizer is composed of two Schottky diodes connected in an anti-parallel configuration as well as a PIN diode in parallel with them. The key point in the design of the PD is the cancellation of the imaginary part of the admittance seen from connection points of the diodes in small-signal operation.

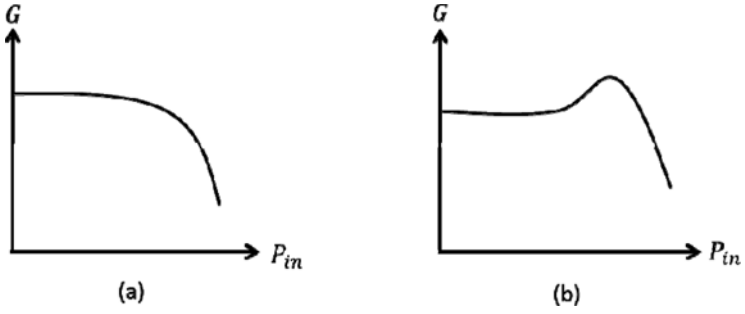

Fig. 5.32 (a) Single and (b) dual inflection amplitude characteristics

Each linearizer generates a nonlinear reflection coefficient, Γ, that relates the input signal, a_1, to the output signal, b_4, using the matrix of the hybrid coupler as follows:

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 1 & -j & 0 \\ 1 & 0 & 0 & -j \\ -j & 0 & 0 & 1 \\ 0 & -j & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ \Gamma \cdot b_2 \\ \Gamma \cdot b_3 \\ 0 \end{bmatrix} \tag{5.77}$$

where $a_2 = \Gamma \cdot b_2$ and $a_3 = \Gamma \cdot b_3$.

Further manipulation of the above matrix results in:

$$b_4 = -\Gamma \cdot a_1 \tag{5.78}$$

Considering the input power $|a_1|^2$ and output power $|b_4|^2$, the following relationship can be derived:

$$P_{out} = |\Gamma|^2 \cdot P_{in} \tag{5.79}$$

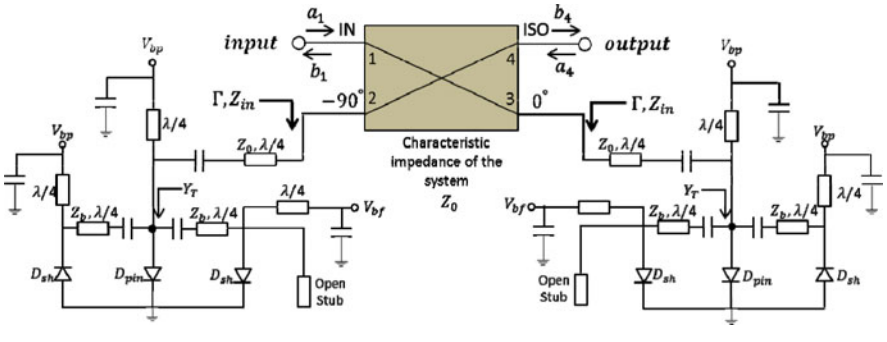

Fig. 5.33 Analog predistorter architecture for linearization of the distortion with dual inflection point

This equation shows that the PD gain is equal to the square value of the reflection coefficient seen from both linearizer branches.

Figure 5.34 shows the performance of the PD in suppressing the distortion of the PA and flattening of the nonlinear part of the AM/AM characteristic of the PA. It is obvious that the distortion is considerably compensated for by the PD.

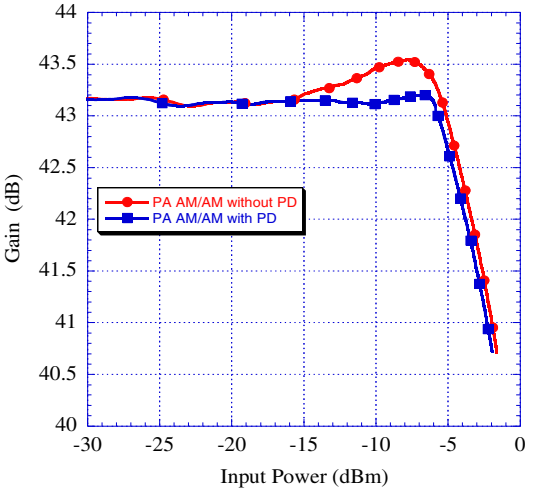

Fig. 5.34 Measured AM/AM of AP603 with PD and without PD versus the input power [23]

5.4.5 Comparison of Linearization Techniques

Table 5.2 compares the performance of feedback, feedforward, and predistortion linearization techniques, in terms of frequency bandwidth, linearity, complexity, power efficiency, and adaptation.

Table 5.2 Performance Comparison of Different Linearization Methods

	Feedback	Feedforward	Analog / Digital Predistortion
Frequency Bandwidth	Narrow	Wide	Ultra/Medium
Linearity	Good	Ultra	Medium/Ultra
Complexity	Medium	High	Medium/Medium
Power Efficiency	High	Low	High/High
Adaptation	Intrinsically adaptive	Intrinsically non-adaptive	Intrinsically non-adaptive

5.5 RF Transmitter Architectures

Power efficient and linear and linear transmitters, such as polar, linear amplification with nonlinear components (LINC), envelope elimination and restoration (EER) and delta-sigma transmitters have received increased attention. These advanced transmitter architectures are theoretically highly efficient and linear, even for sophisticated digital modulation signals with non-constant envelopes and large PAPRs.

5.5.1 Polar Transmitter Architecture

In the polar transmitter, the in-phase (I) and quadrature (Q) baseband signals are transformed from the Cartesian representation to the polar representation. The Cartesian domain signal representation, $S(t) = I(t)\cos\omega_c t + Q(t)\sin\omega_c t$, can be transformed to the polar domain as $S(t) = A(t)\cos(\omega_c t + \varphi(t))$. The transformation can be done with the following equations:

$$A(t) = \sqrt{I^2(t) + Q^2(t)} \tag{5.80}$$

$$\varphi(t) = arct\tan(Q(t)/I(t)) \tag{5.81}$$

A major difference between the Cartesian and polar representations is that the polar basis vectors (amplitude and phase) have widely differing spectral properties compared to the I/Q basis vectors and the modulated RF output. This is clear in (5.80) and (5.81) as $A(t)$ and $\varphi(t)$ are derived from $I(t)$ and $Q(t)$ through nonlinear operations. The nonlinearities in (5.80) and (5.81) cause the polar basis vectors to lose the bandlimited property of the Cartesian I/Q representation. To design the system for high integrity, low EVM and high spectral fidelity, the wideband nature of the amplitude and phase paths has to be considered in the architecture and circuit-level design [24].

In a polar transmitter, the phase, $\varphi(t)$, and amplitude, $A(t)$, components of the signal are processed separately. Figure 5.35 shows a general structure of the polar transmitter. The phase of the carrier is modulated, and it then passes through an amplifier, in which the value of the power supply varies proportionally to the envelope. The PA is shown as a variable-gain amplifier (VGA) in Figure 5.35.

Generally, any signal that modulates the amplitude of the carrier can be employed in a polar transmitter. The phase information extracted from the original signal with a constant or non-constant envelope is converted to a constant envelope signal. This can be achieved by phase modulation with the help of a phase lock loop (PLL) or digital phase lock loop (DPLL) to output the desired transmitting frequencies [25]. The output signal can now be amplified without any concern of distorting the amplitude information.

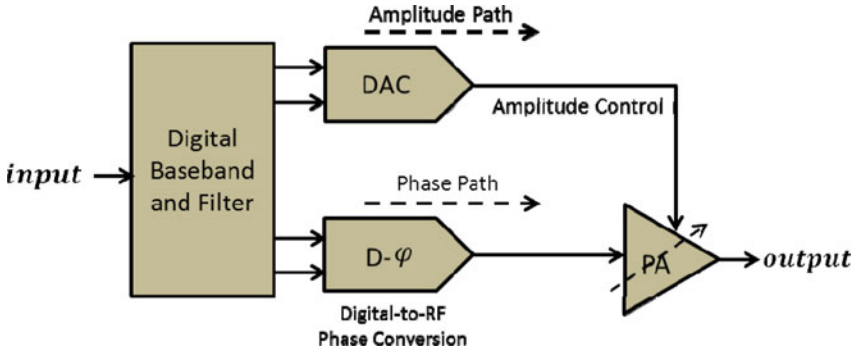

Fig. 5.35 General structure of the polar transmitter

5.5.2 LINC Transmitter Architecture

The principle of the linear amplification with nonlinear components (LINC) amplification system is based on converting the highly varying amplitude modulation of the input signal, $S_{in}(t)$, into phase modulations of two constant envelope signals, $S_1(t)$ and $S_2(t)$, which can be obtained using the following equations:

$$S(t) = r(t) \cdot e^{j \cdot \phi(t)} = S_1(t) + S_2(t) \tag{5.82}$$

$$r(t) = r_{\max} \cdot \cos(\theta(t)) \tag{5.83}$$

hence,

$$\begin{cases} S_1(t) = \dfrac{r_{\max}}{2} \cdot e^{j(\phi(t)+\theta(t))} \\ S_2(t) = \dfrac{r_{\max}}{2} \cdot e^{j(\phi(t)-\theta(t))} \end{cases} \tag{5.84}$$

where $r_{\max}$ represents the maximum of $r(t)$, $\phi(t)$ is the phase of the baseband signal, and $\theta(t)$ is the additional phase modulation angle related to the amplitude of signal, $r(t)$, which is given by:

$$\theta(t) = \arccos\left(\frac{r(t)}{r_{\max}}\right) \tag{5.85}$$

The resultant signal, $S_{out}(t)$, is a linearly amplified version of the input signal, $S_{in}(t)$:

$$S_{out}(t) = \sqrt{2} \cdot G \cdot S_{in}(t) \tag{5.86}$$

where G is the gain of the branch amplifier.

Another way to compute $S_1(t)$ and $S_2(t)$ can illustrated as:

$$\begin{cases} S_1(t) = \dfrac{1}{2} S(t)\left[1 + j \cdot e(t)\right] \\ S_2(t) = \dfrac{1}{2} S(t)\left[1 - j \cdot e(t)\right] \end{cases} \tag{5.87}$$

where $e(t)$ is obtained by:

$$e(t) = \sqrt{\frac{r_{\max}^2}{r^2(t)} - 1} \tag{5.88}$$

Considering that $S_1(t)$ and $S_2(t)$ have constant envelopes, they can be efficiently amplified by means of power efficient or switching-mode nonlinear PAs. The two amplified signals are then combined to retrieve a linearly amplified version of the original amplitude modulated input signal. Therefore, RF PAs can be operated at saturation, which results in maximum power efficiency.

The structure of a LINC amplification system is shown in Figure 5.36. A vector representation of the separated baseband components is also given in Figure 5.37. There are three main components in the LINC transmitter: the signal separator, the nonlinear amplifiers, and the signal combiner [5].

Successful implementation of a LINC transmitter is highly dependent on the precision and control of signal separation, because the LINC architecture is sensitive to amplitude and phase balance. Although there are some analog signal component separation implementations, digital signal component separation implementation offers the flexibility required to apply a control over the precision of the calculations. A digital signal separator also allows implementation of correction algorithms to compensate for any residual imbalance in the analog sections. It is possible to implement the signal separation by implementation of (5.85) or (5.88).

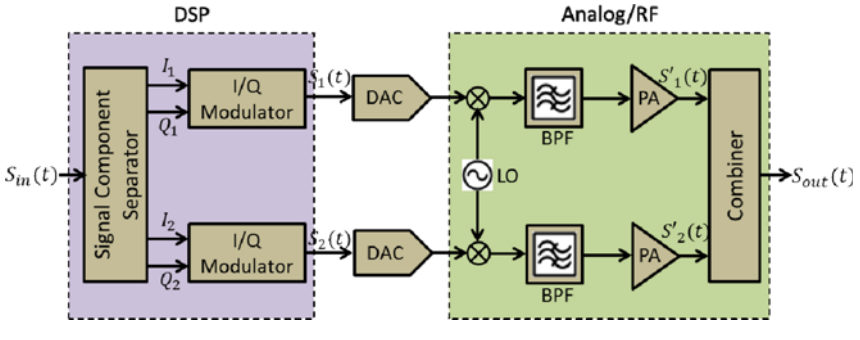

Fig. 5.36 LINC transmitter structure

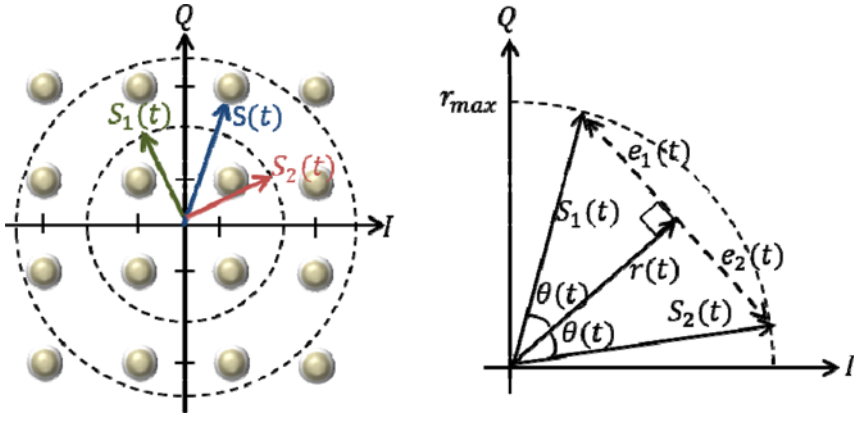

Fig. 5.37 LINC vector decomposition

There are a number of issues that must be addressed in digital implementation of the signal component separation block, such as the use of a look-up table technique, the bit resolution of the computations, the required memory resources, the required computation speed, and the choice of digital-to-analog converters [5].

The PAs should be designed for the highest possible efficiency at saturation, based on the selection of the PA's biasing and impedance matching circuits. For this reason, class F and inverse class F PA designs can be used, which have power efficiencies of around 80%. While the load changes from matched to very high impedance, the PAs should provide a rapid drop-off in DC power consumption in an ideal case.

The bandwidth response of the PA needs to be carefully selected, as the LINC separated signals have larger bandwidth than the original envelope-modulated signal, due to added phase modulation. Additionally, the power amplification block should be made of two identical or quasi-identical amplifiers, in order to preserve the amplitude and phase balance between the two branches [5].

The use of two highly efficient PAs operated with constant envelope signals in the LINC system does not guarantee that the overall efficiency of the LINC transmitter will be high, but the available power at the output of the combiner determines the overall efficiency of the LINC transmitter. In [26], it is shown that the average efficiency of the LINC system also depends on the power distribution function (PDF) of the signal and the type of combiner used. Hence, the combining structure is a key factor in the overall power efficiency and in the linearity of the LINC transmitter. The linearity depends on how the combining structure affects the impedances seen by the amplifiers and how it combines the output signals.

In this framework, the combining structures can be categorized into two classes: 1) matched and isolated combiners, which are also referred to as hybrid or resistive combiners; and, 2) non-matched and non-isolated combiners, which are

also referred to as out-phasing or Chireix combiners [5]. The overall system efficiency, η_{LINC}, of a LINC transmitter using an isolated combiner is given by:

$$\eta_{LINC} = \eta_{PA}^{\max} \cdot \eta_c \tag{5.89}$$

where $\eta_{PA}^{\max}$ is the maximum amplifier efficiency, and η_c is the combiner efficiency. This equation concludes that the LINC system efficiency drops rapidly for low-level signals. The LINC system efficiency can drop as low as $\eta_{PA}^{\max}/10$ at power back-off operation of 10 dB.

In order to make the LINC transmitter applicable for implementation in base stations, some limitations in the LINC transmitter need to be addressed: the gain and phase imbalance between the two branches; and, the bandwidth of the constant envelope signals feeding the PAs. The bandwidths of the phase-modulated signals at the output of the signal separator are generally five times wider than that of the original input signal.

5.5.3 EER Transmitter Architecture

The unmatched efficiency of switching-mode PAs has motivated attempts to use them for linear amplification. Envelope elimination and restoration (EER), proposed by L. R. Kahn in 1952, is one such attempt and is presented in Figure 5.38. A modulated RF signal is split into its polar components, the amplitude signal and the constant amplitude phase-modulated signal by an envelope detector and a limiter, respectively. The limiter output is a constant envelope signal that can be amplified by a highly nonlinear but power efficient PA, usually operating in switching mode, ideally without adding significant AM/AM and AM/PM distortion.

In the envelope path, the amplitude information is extracted by utilizing an envelope detector, which is used to modulate the supply voltage of the PA using the amplitude amplifier / bias modulator. In the phase path, a limiter is utilized to eliminate the amplitude signal variation, which generates the constant amplitude phase-modulated signal to be applied to the PA [6]. The amplitude information is restored to the envelope of the transmitted signal by modulating the supply voltage of the PA, which leads to the EER distinction.

The most significant challenges the EER transmitter faces are mainly related to the bandwidth of the signals to be amplified. The bandwidths of the phase and envelope signals are five times wider than that of the input signal. In fact, both the amplitude and RF power amplifiers should operate over a wide bandwidth, which inevitably reduces their efficiency. The performance of the EER transmitter is significantly affected by the synchronization between the envelope and the RF signal paths [27]. The current state of the art is minimization of the worst-case differential delay between the envelope and RF signal paths.

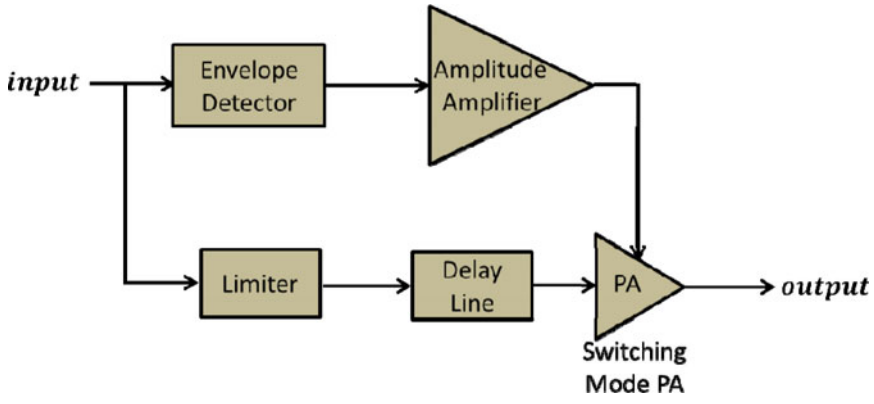

Fig. 5.38 EER transmitter structure

5.5.4 Delta-Sigma Transmitter Architecture

A delta-sigma ($\Delta\Sigma$) modulation based transmitter is another advanced amplification system that transforms the envelope-varying signal to a constant envelope signal and is used with high-efficiency PAs [28]. Figure 5.39 shows the block diagram of an RF delta-sigma based transmitter. Two delta-sigma modulators (DSMs), working at frequency, f_s, are implemented to produce bi-level signals for I and Q signals; and, two high-frequency multiplexers, working at frequency, Nf_s, are used to up-convert the baseband signals to the carrier frequency, $f_c = Nf_s$.

With the use of a third multiplexer, working at $2f_s$, the modulated signals produced by the two multiplexers are combined to generate I and Q signals at carrier frequency f_c. At the output of the third multiplexer, a switching-mode PA is employed to amplify the bi-level I/Q signal. Prior to transmitting the signal through the antenna, a bandpass filter is used to suppress all out-of-band distortion and also to recover the modulated signal around the carrier frequency [1].

The general structure of a DSM is demonstrated in Figure 5.40, in which a quantizer embedded in a loop with a digital-to-analog converter (DAC) in the feedback path is shown. An integrator is in the forward path of the modulator. The input signal to the integrator is the difference between the input signal, $x(t)$, and the quantized output value, $y(t)$, converted back to the predicted analog signal, $\hat{y}(t)$.

Considering that the DAC is ideal and signal delays are negligible, the difference between the input signal, $x(t)$, and the fed back signal, $\hat{y}(t)$, at the integrator input is equal to the quantization error. The quantization noise is denoted by the additive term, $E(t)$. This error is added up in the integrator and then quantized by a 1-bit analog-to-digital converter (ADC).

The output of a DSM is illustrated in the z-domain by:

$$Y(z) = H(z)X(z) + G(z)E(z) \tag{5.90}$$

where $X(z)$, $Y(z)$ and $E(z)$ denote the z-transforms of the input signal, the output signal and the quantization error, respectively.

The signal transfer function, $H(z)$, transforms the signal at the desired frequency band; and, the noise transform function, $G(z)$, suppresses the quantization noise in this band [29].

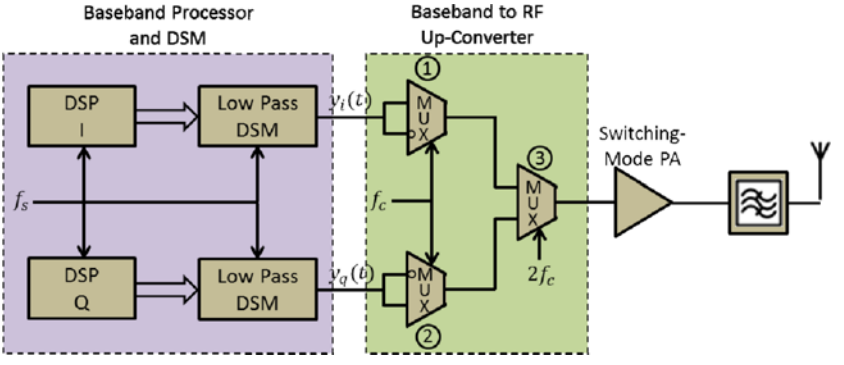

Fig. 5.39 Block diagram of the all-digital delta-sigma transmitter

In the case of a first-order DSM, which is shown in Figure 5.40, the z-domains of the signal and noise transfer functions can be obtained by:

$$H(z) = z^{-1} \tag{5.91}$$

$$G(z) = (1 - z^{-1}) \tag{5.92}$$

There are two DSM-based transmitter architectures in the literature to transform baseband pulsed signals to the desired RF frequency band: the low-pass (LP) DSM-based and the bandpass (BP) DSM-based transmitter architectures [30].

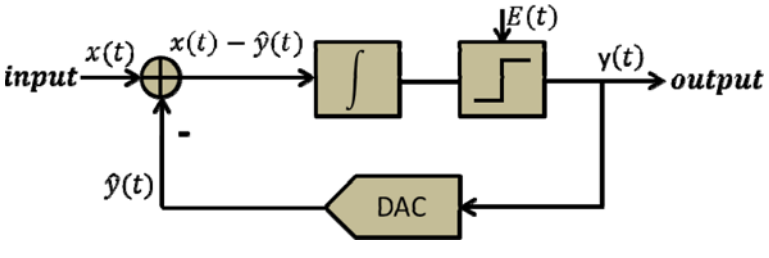

Fig. 5.40 A first-order low-pass DSM general structure

References

[1] Ghannouchi, F.M.: Power Amplifier and Transmitter Architectures for Software Defined Radio Systems. IEEE Circuits and Systems Magazine 10(4), 56–63 (2010)

[2] Larose, C.L., Ghannouchi, F.M.: Optimal Adaptation Methods and Class of Operation: Keys to Improving Feedforward Amplifier Power Efficiency. IEEE Transactions on Vehicular Technology 54(2), 456–467 (2005)

[3] Taijun, L., Bumaiza, S., Ghannouchi, F.M.: Augmented Hammerstein Predistorter for Linearization of Broad-Band Wireless Transmitters. IEEE Transactions on Microwave Theory and Techniques 54(4), 1340–1349 (2006)

[4] Ghannouchi, F.M., Hammi, O.: Behavioural Modeling and Predistortion. IEEE Microwave Magazine 10(7), 52–64 (2009)

[5] Birafane, A., El-Asmar, M., Kouki, A.B., Helaoui, M., Ghannouchi, F.M.: Analyzing LINC Systems. IEEE Microwave Magazine 11(5), 59–71 (2010)

[6] Raab, F.H., Asbeck, P., Cripps, S., Kenington, P.B., Popovic, Z.B., Pothecary, N., Sevic, J.F., Sokal, N.O.: Power Amplifiers and Transmitters for RF and Microwave. IEEE Transactions on Microwave Theory and Techniques 50(3), 814–826 (2002)

[7] Cripps, S.C.: RF Power Amplifiers for Wireless Communications, 2nd edn. Artech House, Norwood (2006)

[8] Kenington, P.B.: High-Linearity RF Amplifier Design. Artech House, Norwood (2000)

[9] Ebrahimi, M.M., Helaoui, M., Ghannouchi, F.M.: Trading-off Stability for Efficiency in Designing Switching-Mode GaN PAs for WiMAX Applications. In: Proc. IEEE Microwave Conference 2009 (APMC 2009), Asia Pacific, pp. 2348–2351 (December 2009)

[10] Grebennikov, A.: RF and Microwave Transmitter Design. John Wiley & Sons (2011)

[11] Hung, T.P.: High Efficiency Switching-Mode Amplifiers for Wireless Communication Systems: ProQuest (2008)

[12] Raab, F.H.: Class-E, Class-C, and Class-F power amplifiers based upon a finite number of harmonics. IEEE Transactions on Microwave Theory and Techniques 49(8), 1462–1468 (2001)

[13] Grebennikov, A., Sokal, N.O.: Switchmode RF Power Amplifiers: Newnes (2007)

[14] Berini, P., Desgagne, M., Ghannouchi, F.M., Bosisio, R.G.: An Experimental Study of the Effects of Harmonic Loading on Microwave MESFET Oscillators and Amplifiers. IEEE Transactions on Microwave Theory and Techniques 42(6), 943–950 (1994)

[15] Raab, F.H.: Maximum Efficiency and Output of Class-F Power Amplifiers. IEEE Transactions on Microwave Theory and Techniques 49(6), 1162–1166 (2011)

[16] Larose, C.L., Ghannouchi, F.M.: Optimization of Feedforward Amplifier Power Efficiency on the Basis of Drive Statistics. IEEE Transactions on Microwave Theory and Techniques 51(1), 41–54 (2003)

[17] Bassam, S.A., Helaoui, M., Ghannouchi, F.M.: Crossover Digital Predistorter for the Compensation of Crosstalk and Nonlinearity in MIMO Transmitters. IEEE Transactions on Microwave Theory and Techniques 57(5), 1119–1128 (2009)

[18] Hammi, O., Ghannouchi, F.M.: Power Alignment of Digital Predistorters for Power Amplifiers Linearity Optimization. IEEE Transactions on Broadcasting 55(1), 109–114 (2009)

[19] Katz, A.: Linearization: Reducing Distortion in Power Amplifiers. IEEE Microwave Magazine 2(4), 37–49 (2001)

[20] Cardinal, J.S., Ghannouchi, F.M.: A New Adaptive Double Envelope Feedback (ADEF) Linearizer for Solid State Power Amplifiers. IEEE Transactions on Microwave Theory and Techniques 43(7), 1508–1515 (1995)

[21] Ghannouchi, F.M.: An S Band RF Digital Linearizer for TWTAs and SSPAs. In: European Conference on Circuit Theory and Design (ECCTD 2009), Antalya, Turkey, pp. 735–738 (August 2009)

[22] Hashmi, M.S., Rogojan, Z.S., Ghannouchi, F.M.: A Flexible Dual-Inflection Point RF Predistortion Linearizer for Microwave Power Amplifiers. Progress in Electromagnetics Research C 13, 1–18 (2010)

[23] Rezaei, S., Hashmi, M.S., Dehlaghi, B., Ghannouchi, F.M.: A Systematic Methodology to Design Analog Predistortion Linearizer for Dual Inflection Power Amplifiers. In: International Microwave Symposium (IMS 2011), Baltimore, Maryland, USA (August 2011)

[24] Nagle, P., Burton, P., Heaney, E., McGrath, F.: A wide-band linear amplitude modulator for polar transmitters based on the concept of interleaving delta modulation. IEEE Journal of Solid-State Circuits 37(12), 1748–1756 (2002)

[25] Staszewski, R.B., Wallberg, J.L., Rezeq, S., Hung, C.M., Eliezer, O.E., Vemulapalli, S.K., Fernando, C., Maggio, K., Staszewski, R., Barton, N., Lee, M.C., Cruise, P., Entezari, M., Muhammad, K., Leipold, D.: All-digital PLL and transmitter for mobile phones. IEEE Journal of Solid-State Circuits 40(12), 2469–2482 (2005)

[26] Birafane, A., Kouki, A.: On the Linearity and Efficiency of Outphasing Microwave Amplifiers. IEEE Transactions on Microwave Theory and Techniques 52(7), 1702–1708 (2004)

[27] Hammi, O., Helaoui, M., Ghannouchi, F.M.: Green Power Amplification Systems for 3G+ Wireless Communication Infrastructure. In: Proc. IEEE Vehicular Technology Conference, pp. 1–5 (September 2010)

[28] Schreier, R., Temes, G.C.: Understanding Delta-Sigma Data Converters. Wiley-IEEE Press (2004)

[29] Helaoui, M., Hatami, S., Negra, R., Ghannouchi, F.M.: A Novel Architecture of Delta-Sigma Modulator Enabling All-Digital Multiband Multistandard RF Transmitters Design. IEEE Transactions on Circuit and Systems II: Express Briefs 55(11), 1129–1133 (2008)

[30] Ghannouchi, F.M., Hatami, S., Aflaki, P., Helaoui, M., Negra, R.: Accurate Power Efficiency Estimation of GHz Wireless Delta-Sigma Transmitters for Different Classes of Switching Mode Power Amplifiers. IEEE Transactions on Microwave Theory and Techniques 58(11), 2812–2819 (2010)