1

Overview

I. A. Glover, S. R. Pennock and P. R. Shepherd

1.1 Introduction

RF and microwave engineering has innumerable applications, from radar (e.g. for air traffic control and meteorology) through electro-heat applications (e.g. in paper manufacture and domestic microwave ovens), to radiometric remote sensing of the environment, continuous process measurements and non-destructive testing. The focus of the courses for which this text was written, however, is microwave communications and so, while much of the material that follows is entirely generic, the selection and presentation of material are conditioned by this application.

Figure 1.1 shows a block diagram of a typical microwave communications transceiver. The transmitter comprises an information source, a baseband signal processing unit, a modulator, some intermediate frequency (IF) filtering and amplification, a stage of up-conversion to the required radio frequency (RF) followed by further filtering, high power amplification (HPA) and an antenna. The baseband signal processing typically includes one, more, or all of the following: an antialising filter, an analogue-to-digital converter (ADC), a source coder, an encryption unit, an error controller, a multiplexer and a pulse shaper. The antialisaing filter and ADC are only required if the information source is analogue such as a speech signal, for example. The modulator impresses the (processed) baseband information onto the IF carrier. (An IF is used because modulation, filtering and amplification are technologically more difficult, and therefore more expensive, at the microwave RF.)

The receiver comprises an antenna, a low noise amplifier (LNA), microwave filtering, a down-converter, IF filtering and amplification, a demodulator/detector and a baseband processing unit. The demodulator may be coherent or incoherent. The signal processing will incorporate demultiplexing, error detection/correction, deciphering, source decoding, digital-to-analogue conversion (DAC), where appropriate, and audio/video amplification and filtering, again where appropriate. If detection is coherent, phase locked loops (PLLs) or their equivalent will feature in the detector design. Other control circuits, e.g. automatic gain control (AGC), may also be present in the receiver.

The various subsystems of Figure 1.1 (and the devices comprising them whether discrete or in microwave integrated circuit form) are typically connected together with transmission

Microwave Devices, Circuits and Subsystems for Communications Engineering Edited by I. A. Glover, S. R. Pennock and P. R. Shepherd

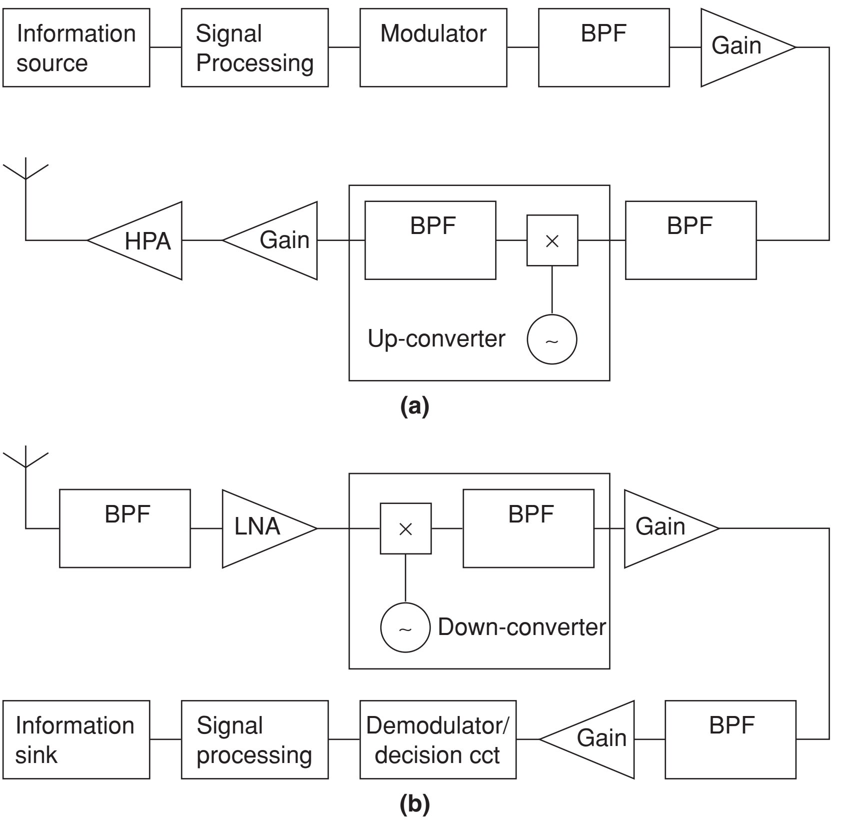

Figure 1.1 Typical microwave communications transmitter (a) and receiver (b)

lines implemented using a variety of possible technologies (e.g. coaxial cable, microstrip, co-planar waveguide).

This text is principally concerned with the operating principles and design of the RF/ microwave subsystems of Figure 1.1, i.e. the amplifiers, filters, mixers, local oscillators and connecting transmission lines. It starts, however, by reviewing the solid-state devices (diodes, transistors, etc.) incorporated in most of these subsystems since, assuming good design, it is the fundamental physics of these devices that typically limits performance.

Sections 1.2–1.7 represent a brief overview of the material in each of the following chapters.

1.2 RF Devices

Chapter 2 begins with a review of semiconductors, their fundamental properties and the features that distinguish them from conductors and insulators. The role of electrons and holes as charge carriers in intrinsic (pure) semiconductors is described and the related concepts of carrier mobility, drift velocity and drift current are presented. Carrier concentration gradients, the diffusion current that results from them and the definition of the diffusion coefficient are also examined and the doping of semiconductors with impurities to increase the concentration of electrons or holes is described. A discussion of the semiconductor energy-band model, which underlies an understanding of semiconductor behaviour, is presented and the important concept of the Fermi energy level is defined. This introductory but fundamental review of semiconductor properties finishes with the definition of mean carrier lifetime and an outline derivation of the carrier continuity equation, which plays a central role in device physics.

Each of the next six major sections deals with a particular type of semiconductor diode. In order of treatment these are (i) simple P-N junctions; (ii) Schottky diodes; (iii) PIN diodes; (iv) step-recovery diodes; (v) Gunn diodes; and (vi) IMPATT diodes. (The use of the term diode in the context of Gunn devices is questionable but almost universal and so we choose here to follow convention.) The treatment of the first three diode types follows the same pattern. The device is first described in thermal equilibrium (i.e. with no externally applied voltage), then under conditions of reverse bias (the P-material being made negative with respect to the N-material), and finally under conditions of forward bias (the P-material being made positive with respect to the N-material). Following discussion of the device's physics under these different conditions, an equivalent circuit model is presented that, to an acceptable engineering approximation, emulates the device's terminal behaviour. It is a device's equivalent circuit model that is used in the design of circuits and subsystems. There is a strong modern trend towards computer-aided design in which case the equivalent circuit models (although of perhaps greater sophistication and accuracy than those presented here) are incorporated in the circuit analysis software. The discussion of each device ends with some comments about its manufacture and a description of some typical applications.

The treatment of the following diode types is less uniform. Step-recovery diodes, being a variation on the basic PIN diode, are described only briefly. The Gunn diode is discussed in some detail since its operating principles are quite different from those of the previous devices. Its important negative resistance property, resulting in its principal application in oscillators and amplifiers, is explained and the relative advantages of its different operating modes are reviewed. Finally, IMPATT diodes are described, that, like Gunn devices, exhibit negative resistance and are used in high power (high frequency) amplifiers and oscillators, their applications being somewhat restricted, however, by their relatively poor noise characteristics. The doping profiles and operating principles of the IMPATT diode are described and the important device equations are presented. The discussion of IMPATT diodes concludes with their equivalent circuit.

Probably the most important solid-state device of all in modern-day electronic engineering is the transistor and it is this device, in several of its high frequency variations, that is addressed next. The treatment of transistors starts with some introductory and general remarks about transistor modelling, in particular, pointing out the difference between small and large signal models. After these introductory remarks three transistor types are addressed in turn, all suitable for RF/microwave applications (to a greater or lesser extent). These are (i) the gallium arsnide metal semiconductor field effect transistor (GaAs MESFET); (ii) the high electron mobility transistor (HEMT); and (iii) the heterojunction bipolar transistor (HBT). In each case the treatment is essentially the same: a short description followed by presentations of the current-voltage characteristic, capacitance-voltage characteristic, the small signal equivalent circuit and the large signal equivalent circuit.

1.3 Signal Transmission and Network Methods

Chapter 3 starts with a survey of practical transmission line structures including those without conductors (dielectric waveguides), those with a single conductor (e.g. conventional waveguide), and those with two conductors (e.g. microstrip). With one exception, all the two-conductor transmission line structures are identified as supporting a quasi-TEM (transverse electromagnetic) mode of propagation – important because this type of propagation can be modelled using classical distributed-circuit transmission line theory. A thorough treatment of this theory is given, starting with the fundamental differential equations containing voltage, current and distributed inductance (L), conductance (G), resistance (R) and capacitance (C), and deriving the resulting line's attenuation constant, phase constant and characteristic impedance. Physical interpretations of the solution of the transmission line equations are given in terms of forward and backward travelling waves and the concepts of loss, dispersion, group velocity and phase velocity are introduced. The frequency-dependent behaviour of a transmission line due to the frequency dependence of its L, G, R and C (due in part to the skin effect) is examined and the special properties of a lossless line (with $R = G = 0$) are derived.

Following the distributed-circuit description of transmission lines, the more rigorous field theory approach to their analysis is outlined. A short revision of fundamental electromagnetic theory is given before this theory is applied to the simplest (TEM) types of transmission line with a uniform dielectric and perfect conductors. The relationship between the time-varying field on the TEM line and the static field solution to Maxwell's equations is discussed and the validity of the solutions derived from this relationship is confirmed. The special characteristics of the TEM propagation mode are examined in some detail. The discussion of basic transmission line theory ends with a physical interpretation of the field solutions and a visualisation of the field distribution in a coaxial line.

Most traditional transmission lines (wire pair, coaxial cable, waveguide) are purchased as standard components and cut to length. Microstrip, and similarly fabricated line technologies, however, are typically more integrated with the active and passive components that they connect and require designing for each particular circuit application. A detailed description of microstrip is therefore given along with the design equations required to obtain the physical dimensions that achieve the desired electrical characteristics, given constraints such as substrate permittivity and thickness that are fixed once a (commercial) substrate has been selected. The limitations of microstrip including dispersion and loss are discussed and methods of evaluating them are presented. The problem of discontinuities is addressed and models for the foreshortened open end (an approximate open circuit termination), vias (an approximate short circuit termination), mitred bends (for reducing reflections at microstrip corners) and T-junctions are described.

In addition to a simple transmission line technology for interconnecting active and passive devices, microstrip can be used as a passive device technology in its own right. Microstrip implementation of low-pass filters is described and illustrated with a specific example. The general theory of coupled microstrip lines, useful for generalised filter and coupler design, is presented and the concepts of odd- and even-modes explained. Equations and design curves for obtaining the physical microstrip dimensions to realise a particular electrical design objective are presented. The directivity of parallel microstrip couplers is discussed and simplified expressions for its calculation are presented. Methods of improving coupler performance by capacitor compensation are described. The discussion of practical microstrip design methods concludes with a brief survey of other microstrip coupler configurations including Lange couplers, branch-line couplers and hybrid rings.

Network methods represent a fundamental way of describing the effect of a device or subsystem inserted between a source and load (which may be the Thévenin/Norton equivalent circuits of a complicated existing system). From a systems engineering perspective, the network parameters of the device or subsystem describe its properties completely – knowledge of the detailed composition of the device/subsystem (i.e. the circuit configuration or values of its component resistors, capacitors, inductors, diodes, transistors, transformers, etc.) being unnecessary. The network parameters may be expressed in a number of different ways, e.g. impedance (z), admittance (y), hybrid (h), transmission line (ABCD) and scattering (s) parameters, but all forms give identical (and complete) information and all forms can be readily transformed into any of the others. Despite being equivalent, there are certain practical advantages and disadvantages associated with each particular parameter set and at RF and microwave frequencies these weigh heavily in favour of using s-parameters. A brief review of all commonly used parameter sets is therefore followed by a more detailed definition and interpretation of s-parameters for both one- and two-port (two- and four-terminal) networks.

The reflection and transmission coefficients at the impedance discontinuities of a device's input and output are described explicitly by the device's s-parameters. One of the central problems in RF and microwave design is impedance matching the input and output of a device or subsystem with respect to its source and load impedances. (This problem may be addressed in the context of a variety of objectives such as minimum reflection, maximum gain or minimum noise figure.) The chapter therefore continues with an account of the most widely used aids to impedance matching, namely the Smith chart and its derivatives (admittance and immitance charts). These aids not only accelerate routine (manual) design calculations but also present a geometrical interpretation of relative impedance that can lead to analytical insights and creative design approaches. Both lumped and distributed element techniques are described including the classic transmission line cases of single and double stub matching. The treatment of matching ends with a discussion of broadband matching, its relationship to quality factor (Q-) circles that can be plotted on the Smith chart, and microstrip line transformers.

Chapter 3 closes with a description of network analysers – arguably the most important single instrument at the disposal of the microwave design engineer. The operating principles of this instrument, which can measure the frequency dependent s-parameters of a device, circuit or subsystem, are described and the sources of measurement error are examined. The critical requirement for good calibration of the instrument is explained and the normal calibration procedures, including the technologies used to make measurements on naked (unpackaged) devices, are presented.

1.4 Amplifiers

Virtually all systems need amplifiers to increase the amplitude and power of a signal. Many people are first introduced to amplifiers by means of low frequency transistor and operational amplifier circuits. At microwave frequencies amplifier design often revolves around terms such as available power, unilateral transducer gain, constant gain and constant noise figure circles, and biasing the transistor through a circuit board track that simply changes its width in order to provide a high isolation connection. This chapter aims to explain these terms and why they are used in the design of microwave amplifiers.

Chapter 4 starts by carefully considering how we define all the power and gain quantities. Microwave frequency amplifiers are often designed using the s-parameters supplied by the device manufacturer, so following the basic definitions of gain, the chapter derives expressions for gain working in terms of s-parameters. These expressions give rise to graphical representations in terms of circles, and the idea of gain circles and their use is discussed.

If we are to realise an amplifier, we want to avoid it becoming an oscillator. Likewise, if we are to make an oscillator, we do not want the circuit to be an amplifier. The stability of a circuit needs to be assessed and proper stability needs to be a design criterion. We look at some basic ideas of stability, and again the resulting conditions have a graphical interpretation as stability circles.

Amplifiers have many different requirements. They might need to be low noise amplifiers in a sensitive receiver, or high power amplifiers in a transmitter. Some applications require narrowband operation while some require broadband operation. This leads to different implementations of microwave amplifier circuits, and some of these are discussed in this chapter.

At microwave frequencies the capacitance, inductance and resistance of the packages holding the devices can have very significant effects, and these features need to be considered when implementing amplifiers in practice. Also alternative circuit layout techniques can be used in place of discrete inductors or capacitors, and some of these are discussed. In dealing with this we see that even relatively simple amplifier circuits are described by a large number of variables. The current method of handling this amount of data and achieving optimum designs is to use a CAD package, and an outline of the use of these is also given.

1.5 Mixers

Mixers are often a key component in a communication or radar system. We generally have our basic message to send, for example, a voice or video signal. This has a particular frequency content that typically extends from very low frequency, maybe zero, up to an upper limit, and we often refer to this as the baseband signal. Many radio stations, TV stations, and mobile phones can be used simultaneously, and they do this by broadcasting their signal on an individually allocated broadcast frequency. It is the mixer circuit that provides the frequency translation from baseband up to the broadcast frequency in the transmitter, and from the broadcast frequency back down to the original baseband in the receiver, to form a superheterodyne system.

A mixer is a non-linear circuit, and must be implemented using a nonlinear component. Chapter 5 first outlines the operation of the commonly used nonlinear components, the diode and the transistor. After that the analysis of these circuits are developed, and the terms used to characterise a mixer are also described. This is done for the so-called linear analysis for small signals, and also for the large signal harmonic balance analysis.

The currently popular transistor mixers are then described, particularly the signal and dual FET implementations that are common in Monolithic Microwave Integrated Circuits (MMICs). The designs of many mixers are discussed and typical performance characteristics are presented. The nonlinear nature of the circuit tends to produce unwanted frequencies at the output of the mixer. These unwanted terms can be 'balanced' out, and the chapter also discusses the operation of single and double balanced mixer configurations.

1.6 Filters

Chapter 6 provides the background and tools for designing filter circuits at microwave frequencies. The chapter begins with a review of two-port circuits and definitions of gain, attenuation and return loss, which are required in the later sections. The various filter characteristics are then described including low-pass, high-pass, band-pass and band-stop responses along with the order number of the filter and how this affects the roll-off of the gain response from the band edges.

The various types of filter response are then described: Butterworth (maximally flat within the passband), Chebyshev (equal ripple response in the pass band), Bessel (maximally flat in phase response) and Elliptic (equal ripple in pass band and stop band amplitude response).

The chapter then addresses the topic of filter realisation and introduces the concept of the low pass prototype filter circuit, which provides the basis for the filter design concepts in the remainder of the chapter. This has a normalised characteristic such that the 3 dB bandwidth of the filter is at a frequency of 1 radian/s and the load impedance is 1 ohm. The four types of filter response mentioned above are then considered in detail, with the mathematical description of the responses given for each.

The chapter then continues with the detail of low pass filter design for any particular value of the order number, N. The equivalent circuit descriptions of the filters are given for both odd and even values of N and also for T- and Π-ladders of capacitors and inductors. Analyses are provided for each of the four filter response types and tables of component values for the normalised response of each is provided.

As these tables are only applicable to the normalised case (bandwidth = 1 radian/s and the filter having a load impedance of 1 ohm), the next stage in the description of the filter realisation is to provide techniques for scaling the component values to apply for any particular load impedance and also for any particular value of low-pass bandwidth. The mathematical relationships between these scalings and the effects on the component values of the filter ladder are derived.

So far, the chapter has only considered the low pass type of filter, so the remainder of the chapter considers the various transformations required to convert the low pass response into equivalent high pass, band pass and band stop responses for each of the filter characteristic types. This therefore provides the reader with all the tools and techniques to design a microwave filter of low pass, high pass, band pass or band stop response with any of the four characteristic responses and of various order number.

1.7 Oscillators and Frequency Synthesisers

Chapter 7 describes the fundamentals of microwave oscillator design, including simple active component realisations using diodes and transistors. The chapter commences with an introduction to solid-state oscillator circuits considered as a device with a load. The fundamental approach is to consider the oscillation condition to be defined so that the sum of the device and load impedances sum to zero. Since the real part of the load impedance must be positive, this implies that the real part of the active device's effective impedance must be negative. This negative resistance is achieved in practice by using a negative resistance diode or a transistor which has a passive feedback network. The active device will have a nonlinear behaviour and its impedance depends on the amplitude of the signal. The balancing condition for the zero impedance condition therefore defines both the frequency and amplitude of oscillation.

The chapter continues with a description of diode realisations of negative resistance oscillators including those based on IMPATT and Gunn devices. This section includes example designs using typical diode characteristics, optimum power conditions and oscillation stability considerations.

Transistor oscillators are then considered. The fundamental design approach is to consider the circuit to be a transistor amplifier with positive feedback, allowing the growth of any starting oscillating signal. This starting signal is most likely to be from ever-present noise in electronic circuits. The feedback circuit is resonant at the desired oscillation frequency, so only noise signals within the bandwidth of the resonant circuit will be amplified and fed back, the other frequencies being filtered out. The possible forms of the resonant feedback circuit are discussed, these include lumped L-C circuits, transmission line equivalents, cavity resonators and dielectric resonators.

The standard topology of transistor feedback oscillators such as the Colpitts, Clapp and Hartley configurations are described and analysed from a mathematical point of view. This section concludes with a discussion of some of the Computer Aided Design (CAD) tools available for the design and analysis of solid-state microwave oscillator circuits.

The next section of the chapter deals with the inclusion of voltage-controlled tuning of the oscillator so that the frequency of oscillation can be varied by the use of a controlling DC voltage. The main implementations for voltage-controlled oscillators (VCOs) use varactor diodes and Yttrium Iron Garnets (YIGs). Varactors are diodes whose junction capacitance can be varied over a significant range of values by the applied bias voltage. When used as part of the frequency selective feedback circuit, variation of this diode capacitance will lead to a variation in the oscillation frequency. YIGs are high-Q resonators in which the ferromagnetic resonance depends (among other factors) on the magnetic field across the device. This in turn can be controlled by an applied voltage. As well as having a high Q value (and therefore a highly stable frequency), YIGs are also capable of a very wide range of voltage control, leading to very broadband voltage control. Examples of practical VCO design complete this section.

The next section considers the characterisation and testing of oscillators. The various parameters used to specify the performance of a particular oscillator include: the oscillator frequency, characterised using a frequency counter or spectrum analyser; the output power, characterised using a power meter; stability and noise (in both amplitude and phase), which can again be analysed using a spectrum analyser or more sophisticated phase noise measurement equipment.

The final section of the chapter deals with phase-locked oscillators, which use phase-locked loops (PLLs) to stabilise the frequency of microwave oscillators. A fundamental description of PLLs is given, along with a consideration of their stability performance. These circuits are then incorporated into the microwave oscillators using a frequency multiplier and harmonic mixers so that the microwave frequency is locked on to a lower, crystal stabilised, frequency so that the characteristics of the highly stable low frequency source are translated on to the microwave frequency.