CHAPTER 12

System noise and communications link budgets

12.1 Introduction

Signal-to-noise ratio (SNR) is a fundamental limiting factor on the performance of communications systems. It can often be improved by a receiving system using appropriate demodulation and signal processing techniques. It is always necessary, however, to know the carrier-to-noise ratio (CNR) present at the input of a communications receiver to enable its performance to be adequately characterised. (The term CNR was used widely in place of SNR in Chapter 11, and is used here, since for most modulation schemes the carrier power is not synonymous with the impressed information signal power.) This chapter reviews some important noise concepts and illustrates how system noise power can be calculated. It also shows how received carrier power can be calculated (at least to first order accuracy) and therefore how a CNR can be estimated.

12.2 Physical aspects of noise

Whilst this text is principally concerned with systems engineering, it is interesting and useful to establish the physical origin of the important noise processes and the physical basis of their observed characteristics (especially noise power spectral density).

Consideration here is restricted to those noise processes arising in the individual components (resistors, transistors, etc.) of a subsystem, often referred to collectively as circuit noise. Noise which arises from external sources and is coupled into a communication system by a receiving antenna is discussed in section 12.4.3.

12.2.1 Thermal noise

Thermal noise is produced by the random motion of free charge carriers (usually electrons) in a resistive medium. These random motions represent small random currents which together produce a random voltage, via Ohm's law, across, for example, the terminals of a resistor. (Such noise does not occur in ideal capacitors as there are no free electrons in a perfect dielectric material.) Thermally excited motion takes place at all temperatures above absolute zero (0 K) since, by definition, temperature is a measure of average kinetic energy per particle. Thermal noise is a limiting factor in the design of many, but not all, radio communications receivers at UHF (0.3 to 3 GHz), Table 1.4, and above.

The power spectral density of thermal noise (at least up to about 10^{12} Hz) can be predicted from classical statistical mechanics. This theory indicates that the average kinetic (i.e. thermal) energy of a molecule in a gas which is in thermal equilibrium and at a temperature T, is $1.5kT$ joules. k is therefore a constant which relates a natural, or atomic, temperature scale to a man-made scale. It is called Boltzmann's constant and has a value of 1.381×10^{-23} J K^{-1}.

The *principle of equipartition* says that a molecule's energy is equally divided between the three dimensions or *degrees of freedom* of space. More generally for a system with any number of degrees of freedom the principle states that:

the average thermal energy per degree of freedom is ½ kT joules.

Equipartition can be applied to the transmission line shown in Figure 12.1. On closing the switches thermal energy will be trapped in the line as standing electromagnetic waves. Since each standing wave mode on the line has two degrees of freedom (its pulsations can exist as sinusoidal or cosinusoidal functions of time) then each standing wave will contain kT joules of energy. The wavelength of the nth mode (Figure 12.2) is given by:

$$\lambda_n = \frac{2l}{n} \quad \text{(m)} \tag{12.1}$$

where l is the length of the transmission line. Therefore the mode number, n, is given by:

$$n = \frac{2lf_n}{c} \tag{12.2}$$

where c is the electromagnetic velocity of propagation on the line. The frequency difference between adjacent modes is:

$$\begin{aligned} f_n - f_{n-1} &= \frac{nc}{2l} - \frac{(n-1)\,c}{2l} \\ &= \frac{c}{2l} \quad \text{(Hz)} \end{aligned} \tag{12.3}$$

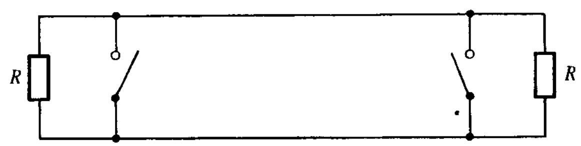

Figure 12.1 *Transmission line for trapping thermal energy.*

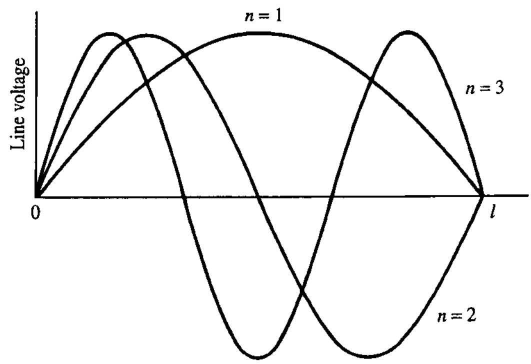

Figure 12.2 *Standing waves of trapped thermal energy on short-circuited transmission line.*

Therefore the number of modes, ν, in a bandwidth B is:

$$\nu = \frac{B}{f_n - f_{n-1}} = \frac{2lB}{c} \tag{12.4}$$

The spatial (line) energy density of each standing wave mode is kT/l (J m^{-1}). Since each standing wave is composed of two travelling waves (with opposite directions of travel) then the spatial energy density in each travelling wave is $\frac{1}{2}kT/l$ (J m^{-1}). The energy per second travelling past a given point in a given direction (i.e. the available power) per oscillating mode is therefore:

$$P_{mode} = \frac{\frac{1}{2}\,kT}{\tau} \quad (\text{J s}^{-1}) \tag{12.5}$$

where τ is the energy transit time from one end of the line to the other. Using $\tau = l/c$:

$$P_{mode} = \frac{\frac{1}{2}\,kTc}{l} \quad (\text{J s}^{-1}) \tag{12.6}$$

The available power, N, in a bandwidth B (i.e. ν modes) is therefore given by:

$$N = P_{mode}\nu \tag{12.7}$$

$$= \frac{\frac{1}{2}\,kTc}{l}\,\frac{2lB}{c} \quad (\text{J s}^{-1})$$

i.e.:

$$N = kTB \quad (\text{W}) \tag{12.8}$$

Equation (12.8) is called Nyquist's formula and is accurate so long as classical mechanics can be assumed to hold. The constant behaviour with frequency, however, of the available (one sided) noise power spectral density, $G_n(f) = kT$, incorrectly predicts infinite power when integrated over all frequencies. This paradox, known as the ultraviolet catastrophe, is resolved when quantum mechanical effects are accounted for.

The complete version of equation (12.8) then becomes:

$$N = \int_0^B G_n(f)\, df \quad \text{(W)} \tag{12.9(a)}$$

where:

$$G_n(f) = \frac{hf}{e^{(hf/kT)} - 1} \quad (\text{W Hz}^{-1}) \tag{12.9(b)}$$

h in equation (12.9(b)) is Planck's constant which has a value of 6.626×10^{-34} (J s). For $hf \ll kT$ then:

$$e^{\left(\frac{hf}{kT}\right)} \approx 1 + \frac{hf}{kT} \tag{12.10}$$

and equation (12.9(b)) reduces to the Nyquist form:

$$G_n(f) \approx kT \quad (\text{W Hz}^{-1}) \tag{12.11}$$

The classical and quantum mechanical power spectral densities of thermal noise are compared in Figure 12.3.

The Thévenin equivalent circuit of a thermal noise source is shown in Figure 12.4. Since the available noise power for almost all practical temperatures and frequencies (excluding optical applications) is kTB W then the maximum available RMS noise voltage V_{na} across a (conjugately) matched load is:

$$V_{na} = \sqrt{NR} = \sqrt{kTBR} \quad \text{(V)} \tag{12.12}$$

where R is the load (and source) resistance. Since the same voltage is dropped across the

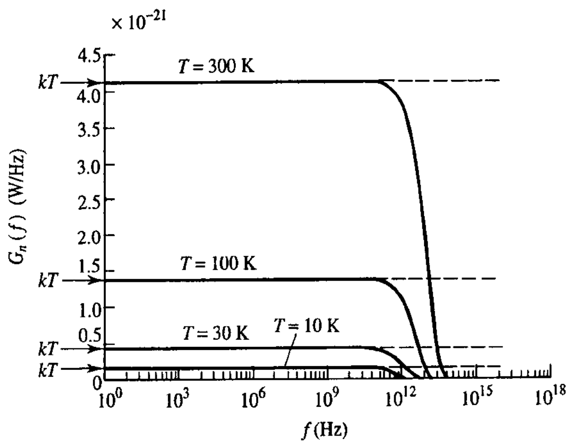

Figure 12.3 *Power spectral densities of thermal noise predicted by classical (- - -) and quantum (---) mechanics at four temperatures.*

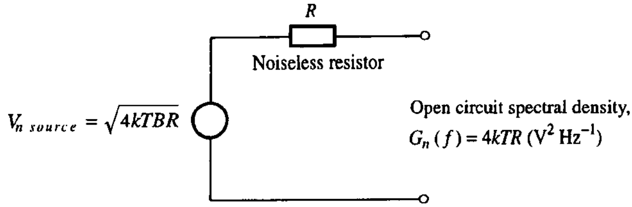

Figure 12.4 *Thévenin equivalent circuit of a thermal noise source.*

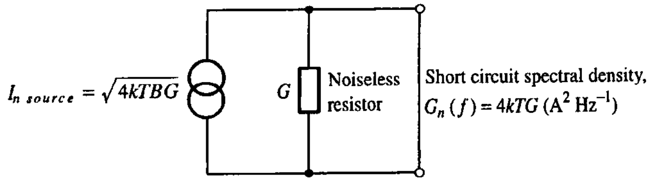

Figure 12.5 *Norton equivalent circuit of a thermal noise source.*

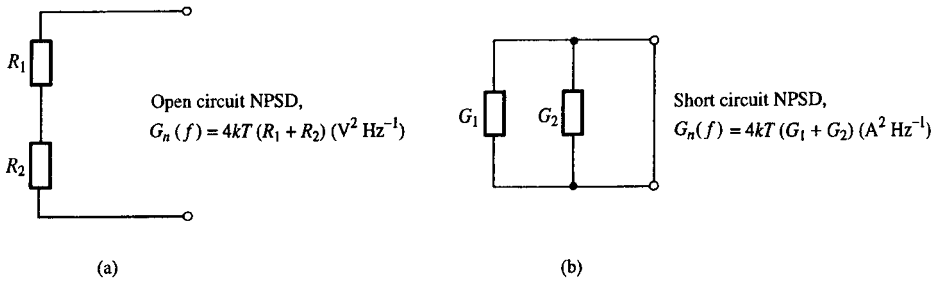

Figure 12.6 *NPSD for resistors in (a) series and (b) parallel.*

source and load resistances the equivalent RMS noise voltage of the source (i.e. the open circuit voltage of the Thévenin equivalent circuit) is given by:

$$V_{n\ source} = 2\sqrt{kTBR} = \sqrt{4kTBR} \quad \text{(V)} \tag{12.13}$$

An equivalent random current source can also be defined using the Norton equivalent circuit in Figure 12.5. In this case the current source has a RMS value given by:

$$I_{n\ source} = \sqrt{4kTBG} \quad \text{(A)} \tag{12.14}$$

where $G = 1/R$ is the load (and source) conductance. For passive components (such as resistors) connected in series and parallel their equivalent mean square noise voltages and currents add in an obvious way. For example, the series combination of resistances R_1 and R_2 in Figure 12.6(a) results in:

$$V^2_{n\ source} = V^2_{n\ source\ 1} + V^2_{n\ source\ 2} \quad (\text{V}^2) \tag{12.15(a)}$$

i.e.:

$$V_{n\,source} = \sqrt{4kTB(R_1 + R_2)} \quad \text{(V)} \tag{12.15(b)}$$

and the parallel combination of conductances G_1 and G_2 in Figure 12.6(b) results in:

$$I^2_{n\,source} = I^2_{n\,source\,1} + I^2_{n\,source\,2} \quad (\text{A}^2) \tag{12.16(a)}$$

i.e.:

$$I_{n\,source} = \sqrt{4kTB(G_1 + G_2)} \quad \text{(A)} \tag{12.16(b)}$$

Alternatively, and entirely unsurprisingly, the equivalent voltage source of the parallel combination is given by:

$$V_{n\,source} = \frac{I_{n\,source}}{G}$$

$$= \sqrt{4kTB\,\frac{R_1\,R_2}{R_1 + R_2}} \quad \text{(V)} \tag{12.16(c)}$$

i.e. the two parallel resistors together behave as a single resistor with the appropriate parallel value.

12.2.2 Non-thermal noise

Although the time averaged current flowing in a device may be constant, statistical fluctuations will be present if individual charge carriers have to pass through a potential barrier. The potential barrier may, for example, be the junction of a pn diode, the cathode of a vacuum tube or the emitter-base junction of a bipolar transistor. Such statistical fluctuations constitute shot noise. The traditional device used to illustrate the origin of shot noise is a vacuum diode, Figure 12.7. Electrons are emitted thermally from the cathode. Assuming there is no significant space charge close to the cathode surface due to previously emitted electrons (i.e. assuming that the diode current is temperature limited), then electrons are emitted according to a Poisson statistical process (see Chapter 17).

The spatial (line) charge density, σ, between cathode and anode is given by:

$$\sigma = \frac{nq_e}{l} \quad \text{(C/m)} \tag{12.17}$$

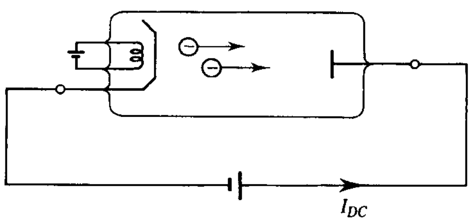

Figure 12.7 *Schematic diagram for a vacuum diode.*

where n is the average number of electrons in flight, l is the distance between cathode and anode and q_e is the charge on an electron. If the average electron velocity is v then the average current, I_{DC}, in (any part of) the circuit is:

$$I_{DC} = \sigma v \quad \text{(C/s)} \tag{12.18}$$

and the average contribution to I_{DC} from any individual electron is:

$$i_e = \frac{I_{DC}}{n} = q_e \frac{v}{l} \quad \text{(C/s)} \tag{12.19}$$

The current associated with any given electron can therefore be considered to flow only for that time the electron is in flight. Since l/v is the average cathode–anode transit time, τ, for the electrons, equation (12.19) can be written as:

$$i_e = \frac{q_e}{\tau} \quad \text{(A)} \tag{12.20}$$

Thus each electron gives rise to a current pulse of duration τ as it moves from cathode to anode. (The precise shape of the current pulse will depend on the way in which the electron's velocity varies during flight. For the purpose of this discussion, however, the pulse can be assumed to be rectangular, Figure 12.8.)

The total current will be a superposition of such current pulses occurring randomly in time with Poisson statistics. Since each pulse has a spectrum with (nominal) bandwidth $1/\tau$ Hz (in the ideal rectangular pulse case the energy spectrum will have a sinc^2 shape, Figure 12.9) then for frequencies very much less than $1/\tau$ (Hz) shot noise has an essentially white spectrum.

The (two sided) current spectral density of the pulse centered on $t = 0$ due to a single electron is:

$$I(f) = \int_{-\infty}^{\infty} \frac{q_e}{\tau} \Pi\left(\frac{t}{\tau}\right) e^{-j2\pi ft}\, dt$$

$$= q_e \,\text{sinc}(\tau f) \quad \text{(A/Hz)} \tag{12.21}$$

The (normalised, two sided) energy spectral density of this current pulse is therefore:

$$E_i(f) = q_e^2 \,\text{sinc}^2(\tau f) \quad (\text{A}^2\ \text{s/Hz}) \tag{12.22}$$

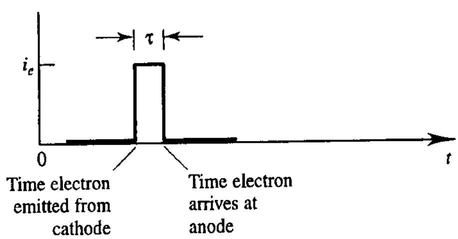

Figure 12.8 *Current pulse due to a single electron in diode circuit of Figure 12.7.*

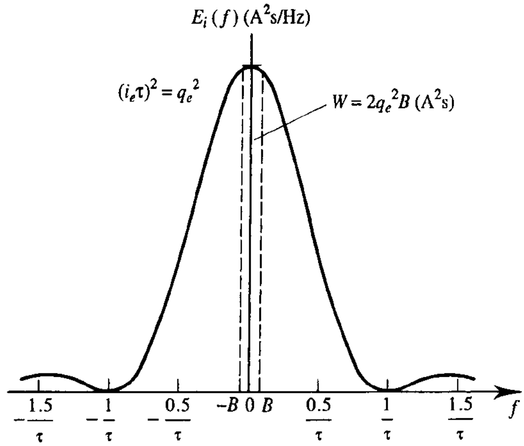

Figure 12.9 *Energy spectral density of current pulse shown in Figure 12.8.*

For frequencies much less than $1/\tau$ Hz, $E_i(f)$ is a constant, namely q_e^2.

Using this model the energy, W, contained within a bandwidth B ($\ll 1/\tau$) is given by:

$$W = 2q_e^2 B \quad (\text{A}^2\ \text{s}) \tag{12.23}$$

For m electrons arriving at the anode in a time T the total shot noise (normalised) power will therefore be:

$$I_n^2 = \frac{2mq_e^2 B}{T} \quad (\text{A}^2) \tag{12.24}$$

mq_e/T, however, can be identified as the total DC current flowing in the circuit, i.e.:

$$I_n = \sqrt{2I_{DC}q_e B} \quad (\text{A}) \tag{12.25}$$

Equation (12.25) shows that shot noise RMS current is proportional to both the square root of the DC current and the square root of the measurement bandwidth. Although this result has been derived for a vacuum diode it is also correct for a pn junction diode and for the base and collector (with $I_{DC} = I_B$ and $I_{DC} = I_C$ as appropriate) of a bipolar junction transistor. Figure 12.10 shows a bipolar transistor along with its equivalent thermal and shot noise sources. $G_{V_{nB}}$ is the power spectral density of the thermal noise voltage generated in the transistor's base spreading resistance. $G_{I_{nB}}$ is the power spectral density of the shot noise current associated with those carriers flowing across the emitter-base junction which subsequently arrive at the base terminal of the transistor. (The fluctuation in base current caused by $G_{V_{nB}}$ and $G_{I_{nB}}$ will be amplified, like any other base current changes, at the transistor collector.) $G_{I_{nC}}$ is the power spectral density of the shot noise current associated with the carriers flowing across the emitter-base junction which subsequently arrive at the transistor collector. For FETs the shot noise is principally associated with gate current and can be modelled by equation (12.25) with $I_{DC} = I_G$.

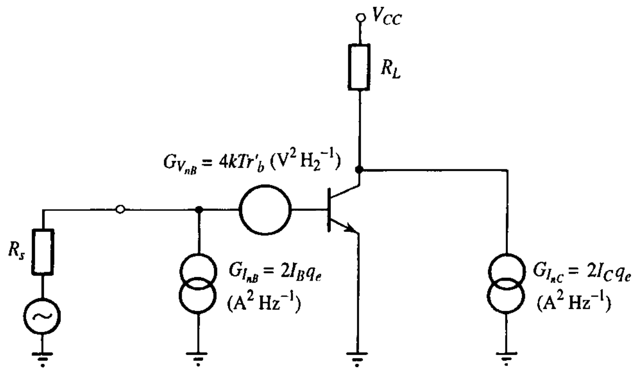

Figure 12.10 *Power spectral densities of thermal (G_{V_n}) and shot (G_{I_n}) noise processes in a bipolar transistor.*

Flicker, or $1/f$, noise also occurs in most active (and some passive) devices. It is rather device specific and is therefore not easily modelled. Its power spectral density has approximately a $1/f$ characteristic for frequencies below a few kilohertz. Above a few kilohertz flicker noise has an essentially flat spectrum but in any case is weak and usually neglected. Because flicker noise power is concentrated at low frequencies it is sometimes called pink noise.

12.2.3 Combining white noise sources

In many applications it is sufficient to account for thermal noise and shot noise only. This also has the practical advantage that circuit noise calculations can be made initially on a power spectral density basis, the total power then being found using the circuit noise bandwidth.

Consider Figure 12.11(a). There are two principal sources of shot noise and, effectively, three sources of thermal noise in this circuit. The former arise from the base-emitter junction of the transistor and the latter arise from the source resistance, R_s (comprising the ouput resistance of the preceding stage in parallel with the biasing resistors R_1 and R_2), the base spreading resistance, $r_{b'}$, and the load resistance, R_L. Figure 12.11(b) shows an equivalent circuit of the transistor, incorporating its noise sources, and Figure 12.11(c) shows an equivalent circuit of Figure 12.11(a) with all noise sources (excluding that due to the load) transferred to the transistor's input and represented as RMS noise voltages. The noise sources in Figure 12.11(c) are given by:

$$V_{nr_{b'}}^2 = 4\,kTr_{b'}B \tag{12.26(a)}$$

$$V_{nB}^2 = I_{nB}^2\,(R_s + r_{b'})^2 = 2q_e\,I_B\,B\,(R_s + r_{b'})^2 \tag{12.26(b)}$$

$$\begin{aligned} V_{nC}^2 &= (I_{nC}/h_{fe})^2\,(R_s + r_{b'} + r_\pi)^2 \\ &= (2q_e\,I_C\,B)\,(R_s + r_{b'} + r_\pi)^2/(g_m\,r_\pi)^2 \end{aligned} \tag{12.26(c)}$$

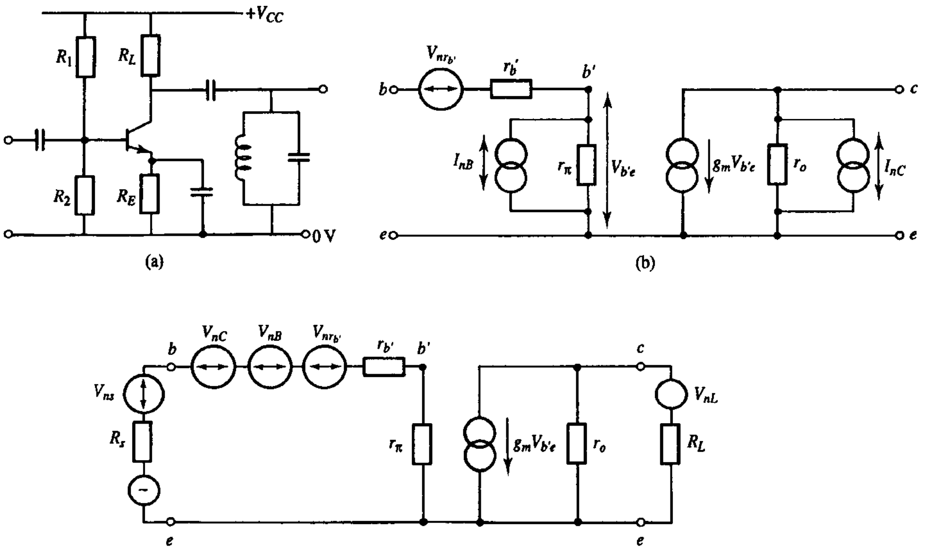

Figure 12.11 *(a) Single stage transistor amplifier; (b) transistor equivalent circuit with noise sources; and (c) equivalent circuit including driving, biasing and load components with most noise sources referred to input.*

$$V_{ns}^2 = 4\,kTR_sB \tag{12.26(d)}$$

$$V_{nL}^2 = 4\,kTR_LB \tag{12.26(e)}$$

where B is the noise bandwidth of the circuit, q_e is the electronic charge, I_B and I_C are the quiescent base and collector currents respectively, $h_{fe} = g_m r_\pi$ is the small signal forward current gain of the transistor and g_m is the transistor's transconductance.

The, effective, transistor input noise (which excludes that due to the source resistance) is therefore given by:

$$\begin{aligned}V_{n\,tran}^2 &= V_{nC}^2 + V_{nB}^2 + V_{nr_{b'}}^2 \\ &= (2q_eI_CB)(R_s + r_{b'} + r_\pi)^2/(g_mr_\pi)^2 + 2q_eI_BB(R_s + r_{b'})^2 + 4\,kTr_{b'}B \\ &= 4\,kT\left[0.5(R_s + r_{b'} + r_\pi)^2/(g_mr_\pi^2) + 0.5g_m(R_s + r_{b'})^2/h_{FE} + r_{b'}\right]B \\ &= 4\,kTR_{tran}B \end{aligned} \tag{12.27}$$

where h_{FE} is the transistor's DC forward current gain and use has been made of $g_m = q_eI_C/(kT) \approx 40I_C$ at room temperature. R_{tran}, defined by the square bracket in equation (12.27), represents a hypothetical resistor located at the (noiseless) transistor's input which would account (thermally) for the actual noise introduced by the real (noisy) transistor. The total noise at the transistor output (including that, V_{nL}^2, generated by the load resistor, R_L) is therefore:

$$V_{n\,o/p}^2 = (V_{ns}^2 + V_{n\,tran}^2)\left(\frac{r_\pi}{R_s + r_{b'} + r_\pi}\right)^2 g_m^2 \left(\frac{r_o\, R_L}{r_o + R_L}\right)^2 + V_{nL}^2 \tag{12.28}$$

Since the source resistance and load resistance depend, at least partly, on the output impedance of the preceding stage and input impedance of the following stage the importance of impedance optimisation in low noise circuit design can be appreciated. The following example illustrates a circuit noise calculation and also shows, explicitly, how such a calculation can be related to the conventional figure of merit (i.e. noise factor or noise figure, see section 12.3.3) commonly used to summarise the noise performance of a circuit, system or subsystem. (In practice communications circuits usually operate at high frequency where s-parameter transistor descriptions are appropriate. For a full treatment of low noise, high frequency, circuit design techniques the reader is referred to specialist texts, e.g. [Smith, Yip, Liao].)

EXAMPLE 12.1

The transistor in figure 12.11(a)) has the following parameters: $r_{b'} = 150\ \Omega$, $r_o = 40\ \text{k}\Omega$, $h_{fe} = 100$, $h_{FE} = 80$. The circuit in which the transistor is embedded has the component values: $R_1 = 10\ \text{k}\Omega$, $R_2 = 3.3\ \text{k}\Omega$, $R_E = 1\ \text{k}\Omega$, $R_L = 3.3\ \text{k}\Omega$. The output resistance of the signal source which drives this circuit is 1.7 kΩ. Find: (a) the transistor's equivalent thermal input noise resistance, R_{tran}; (b) the noise power spectral density at the transistor outputs; and (c) the ratio of source plus transistor noise to transistor noise only. (For part (b) assume that, apart from thermal noise arising from the signal source's output resistance, the source is noiseless.) In part (c) the ratio which you have been asked to calculate is a noise factor, see section 12.3.3.

(a) The transistor transconductance, g_m, depends on the collector current which is found in the usual way, i.e.:

$$R_{in} \approx (h_{FE} + 1)\, R_E = (80 + 1)\, 1 \times 10^3 = 81\ \text{k}\Omega$$

$$V_B = \frac{R_2 R_{in} / (R_2 + R_{in})}{[R_2 R_{in} / (R_2 + R_{in})] + R_1}\, V_{CC}$$

$$= \frac{(3.3 \times 81) / (3.3 + 81)}{[(3.3 \times 81) / (3.3 + 81)] + 10} \times 10 = 2.4\ \text{V}$$

$$V_E = V_B - V_{BE} = 2.4 - 0.6 = 1.8\ \text{V}$$

$$I_C \approx I_E = \frac{V_E}{R_E} = \frac{1.8}{1 \times 10^3} = 1.8\ \text{mA}$$

The transconductance is the reciprocal of the intrinsic emitter resistance, r_e, i.e.:

$$g_m = \frac{1}{r_e} = \frac{I_C\ (\text{mA})}{25} = \frac{1.8}{25} = 0.072\ \text{S}$$

$$r_\pi = \frac{h_{fe}}{g_m} = \frac{100}{0.072} = 1400\ \Omega$$

The source resistance, R_s, is the parallel combination of the signal source's output resistance, R_o, and the transistor's biasing resistors:

$$R_s = \left[\frac{1}{R_o} + \frac{1}{R_1} + \frac{1}{R_2} \right]^{-1}$$

$$= \left[\frac{1}{1700} + \frac{1}{10000} + \frac{1}{3300} \right]^{-1} = 1000 \ \Omega$$

From equation (12.27):

$$R_{tran} = \frac{0.5\,(R_s + r_{b'} + r_\pi)^2}{g_m\, r_\pi^2} + \frac{0.5\, g_m\,(R_s + r_{b'})^2}{h_{FE}} + r_{b'}$$

$$= \frac{0.5\,(1000 + 150 + 1400)^2}{(0.072)\,(1400^2)} + \frac{0.5\,(0.072)\,(1000 + 150)^2}{80} + 150$$

$$= 23 + 600 + 150 = 770\ \Omega$$

(b) The total noise power expected at the transistor output is the sum of contributions from the source resistance, the transistor and the load. Thus, using equation (12.28):

$$V_{n\,o/p}^2 = (V_{ns}^2 + V_{n\,tran}^2)\left(\frac{r_\pi}{R_s + r_{b'} + r_\pi}\right)^2 g_m^2 \left(\frac{r_o\,R_L}{r_o + R_L}\right)^2 + V_{nL}^2$$

where $V_{ns}^2 = 4\ kTR_sB$ and $V_{nL}^2 = 4\ kTR_LB$. Therefore:

$$V_{n\,o/p}^2 = 4\ kT\left[(R_s + R_{tran})\left(\frac{r_\pi}{R_s + r_{b'} + r_\pi}\right)^2 g_m^2 \left(\frac{r_o\,R_L}{r_o + R_L}\right)^2 + R_L\right] B\ \ (\mathrm{V}^2)$$

$$= 4 \times 1.38 \times 10^{-23} \times 290 \left[(1000 + 770)\left(\frac{1400}{1150 + 1400}\right)^2 0.072^2 \left(\frac{40000 \times 3300}{40000 + 3300}\right)^2 + 3300\right] B$$

$$G_{n\,o/p} = V_{n\,o/p}^2 \,/\, B$$

$$= 4 \times 4 \times 10^{-21}\ \ [(1000 + 770)\ 15000 + 3300]$$

$$= 16 \times 10^{-21}\left[27 \times 10^6 + 3.3 \times 10^3\right] = 4.3 \times 10^{-13}\ \ \mathrm{V^2/Hz}$$

(c) Noise factor, f, (see section 12.3.3) is given by:

$$f = \frac{R_s + R_{tran}}{R_s} = \frac{1000 + 770}{1000} = 1.77$$

12.3 System noise calculations

The gain, G, of a device is often expressed in decibels, as the ratio of the output to input voltages or powers, i.e., $G_{dB} = 20\log_{10}(V_o/V_i)$ or $G_{dB} = 10\log_{10}(P_o/P_i)$. Being a ratio gain is not related to any particular power level. There is a logarithmic unit of power, however, which defines power in dB above a specified reference level. If the reference power is 1 mW the units are denoted by dBm (dB with respect to a 1 mW reference) and if the reference level is 1 W the units are denoted by dBW. 10 mW would thus correspond to +10 dBm and 100 mW would be +20 dBm, etc. dBm can be converted to dBW using +30 dBm ≡ 0 dBW. Table 12.1 illustrates these relationships.

The ways in which the noise performance of individual subsystems is specified, and the way these specifications are used to calculate the overall noise performance of a complete system, are now discussed. The overall noise characteristics of individual subsystems are usually either specified by the manufacturer or measured using special instruments (e.g. noise figure meters). If $T = 290$ K (ambient or room temperature) and $B = 1$ Hz, then the available noise power is $1.38 \times 10^{-23} \times 290 \times 1 = 4 \times 10^{-21}$ W/Hz. Expressing this power (in a 1 Hz bandwidth) in dBW gives a noise power spectral density of −204 dBW/Hz or −174 dBm/Hz. Power, in dBm or dBW, can be scaled by simply adding or subtracting the gain of amplifiers or attenuators, measured in dB, to give directly the output power, again in dBm or dBW.

Table 12.1 *Signal power measures.*

dBW	*Power level* (W)	dBm
30 dBW	1,000	60 dBm
20 dBW	100	50 dBm
10 dBW	10	40 dBm
0 dBW	1	30 dBm
−10 dBW	1/10	20 dBm
−20 dBW	1/100	10 dBm
−30 dBW	1/1,000	0 dBm
−40 dBW	1/10,000	−10 dBm

12.3.1 Noise temperature

Consider again equation (12.8) which defines available noise power. The actual noise power delivered is less than this if the source and load impedances are not matched (in the maximum power transfer, i.e. conjugate impedance, sense). A convenient way to specify noise power is via an equivalent thermal noise temperature, T_e, given by:

$$T_e = \frac{N}{kB} \quad \text{(K)} \tag{12.29}$$

The noise temperature of a subsystem (say an amplifier) is *not* the temperature of the room it is in *nor* even the temperature inside its case. It is the (hypothetical) temperature which an ideal resistor matched to the input of the subsystem would need to be at in order to account for the extra available noise observed at the device's output *over and above that which is due to the actual input noise.* This idea is illustrated schematically in Figure 12.12. The total available output noise of a device is therefore given by:

$$N = (kT_s B + kT_e B)G \tag{12.30}$$

where the first term in the brackets is the contribution from the input (or source) noise and the second term is the contribution from the subsystem itself. G is the (power) gain

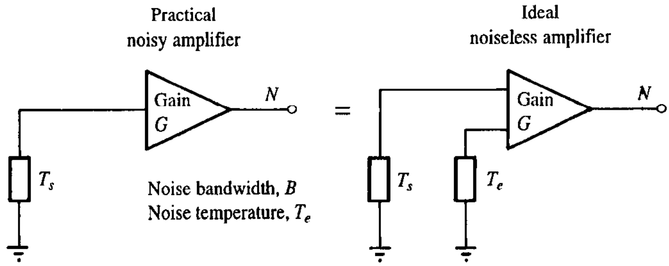

Figure 12.12 *Equivalent noise temperature (T_e) of an amplifier.*

(expressed as a ratio) of the subsystem which can be greater or less than 1.0. For example, an amplifier would have a gain exceeding unity while the mixer in a downconverter, Figure 1.4, would typically have a gain of approximately 0.25. Whilst the contribution of the subsystem to the total output noise, and therefore the noise temperature, will generally be *influenced* by its physical temperature, it will also depend on its design and the quality of its component parts, particularly with respect to the sections at, or close to, its 'front' (i.e. input) end.

Some non-thermal noise processes, such as $1/f$ or flicker noise, have non-white spectra which if included in equation (12.29) would make T_e a function of bandwidth. In practice, however, for most systems work, white noise is assumed to dominate, making T_e bandwidth independent. If thermal noise is dominant (as is often the case in low noise systems) then physical cooling of the device will improve its noise performance.

It has been stated above that the equivalent noise temperature of a device has a dependence on its physical temperature. The strength of this dependence is determined by the relative proportions of thermal to non-thermal noise which the device generates. Strictly speaking, therefore, the noise temperature quoted for a device or subsystem relates to a specific device physical temperature. This temperature is invariably the equilibrium temperature which the device attains (under normal operating conditions which may incorporate heat sinks etc.) when the ambient temperature around it is 290 K. The operating environments of electronic devices normally have ambient temperatures which are sufficiently close to this to make errors in noise calculations based on an assumed temperature of 290 K negligible for most practical purposes.

EXAMPLE 12.2

Calculate the output noise of the amplifier, shown in Figure 12.12, assuming that the amplifier gain is 20 dB, its noise bandwidth is 1 MHz, its equivalent noise temperature is 580 K and the noise temperature of its source is 290 K.

The available NPSD from the matched source at temperature T_0 is given in equation (12.11) by:

$$G_{n,s}(f) = kT_0 \quad \text{(W/Hz)}$$

(T_0 is a widely used symbol denoting 290 K.) The available NPSD from the equivalent resistor at temperature T_e (modelling the internally generated noise) is:

$$G_{n,e}(f) = kT_e \quad \text{(W/Hz)}$$

The total equivalent noise power at the input is:

$$N_{in} = k(T_0 + T_e)B \quad \text{(W)}$$

The total noise power at the amplifier output is therefore:

$$\begin{aligned} N_{out} &= Gk(T_0 + T_e)B \quad \text{(W)} \\ &= 10^{20/10} \times 1.38 \times 10^{-23} \times (290 + 580) \times 1 \times 10^6 \\ &= 1.2 \times 10^{-12} \text{ W} = -119.2 \text{ dBW or } -89.2 \text{ dBm} \end{aligned}$$

12.3.2 Noise temperature of cascaded subsystems

The total system noise temperature, at the output of a device or subsystem, $T_{syst\ out}$, can be found by dividing equation (12.30) by kB, i.e.:

$$T_{syst\ out} = (T_s + T_e)G \tag{12.31}$$

If several subsystems are cascaded, as shown in Figure 12.13, the noise temperatures at the output of each subsystem are given by:

$$T_{out\ 1} = (T_s + T_{e1})G_1 \tag{12.32(a)}$$

$$T_{out\ 2} = (T_{out\ 1} + T_{e2})G_2 \tag{12.32(b)}$$

$$T_{out\ 3} = (T_{out\ 2} + T_{e3})G_3 \tag{12.32(c)}$$

The noise temperature at the output of subsystem 3 is therefore:

$$\begin{aligned} T_{out\ 3} &= \{[(T_s + T_{e1})G_1 + T_{e2}]G_2 + T_{e3}\}G_3 \qquad (12.33) \\ &= T_sG_1G_2G_3 + T_{e1}G_1G_2G_3 + T_{e2}G_2G_3 + T_{e3}G_3 \end{aligned}$$

This temperature can be referred to the input of subsystem 1 by dividing equation (12.33)

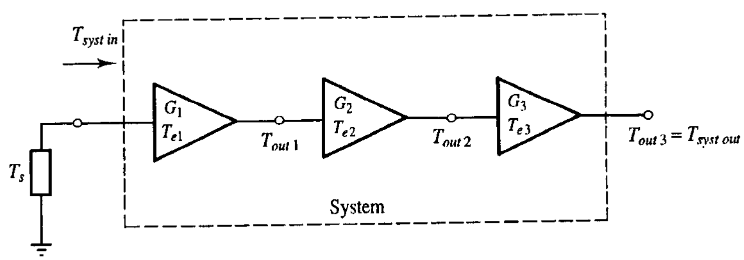

Figure 12.13 *Partitioning of system into cascaded amplifier subsystems.*

by the total gain, $G_1G_2G_3$, i.e.:

$$T_{syst\ in} = T_s + T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1G_2} \tag{12.34}$$

or:

$$T_{syst\ in} = T_s + T_e \tag{12.35}$$

where T_e is the equivalent (input) noise temperature of the system excluding the source and $T_{syst\ in}$ is the overall input noise temperature *including the source*. The total noise temperature output of a set of cascaded subsystems is then simply given by:

$$T_{syst\ out} = T_{syst\ in}G_1G_2G_3 \tag{12.36}$$

It is important to remember that the 'gain' of the individual subsystems can be greater *or less* than 1.0. In the latter case the subsystem is lossy. (A transmission line, for instance, with 10% power loss would have a gain of 0.9 or −0.46 dB.) The equivalent (input) noise temperature, $T_{e,l}$, of a lossy device at a physical temperature T_{ph}, can be found from its 'gain' using:

$$T_{e,l} = \frac{T_{ph}(1 - G_l)}{G_l} \tag{12.37}$$

(The subscript l here reminds us that a lossy device is being considered and that the gain G_l is therefore less than 1.0.) Equation (12.37) is easily verified by considering a transmission line terminated in matched loads as shown in Figure 12.14. Provided the loads, R_s and R_L, and the transmission line are in thermal equilibrium there can be no net flow of noise power across the terminals at either end of the transmission line. From Nyquist's noise formula, equation (12.8), the power available to the transmission line from R_s is:

$$N_s = kT_{ph}B \tag{12.38}$$

If the transmission line has (power) gain, G_l (< 1.0), then the power available to R_L *from* R_s is:

$$N_s' = kT_{ph}BG_l \tag{12.39}$$

Since R_L supplies

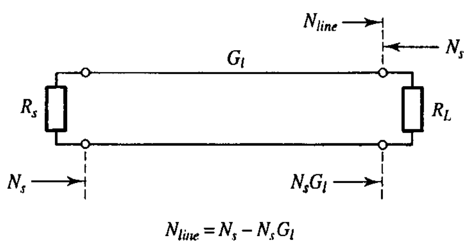

Figure 12.14 *Source, transmission line and load in thermal equilibrium.*

$$N_L = kT_{ph}B \tag{12.40}$$

of power to the transmission line then the transmission line itself must supply the balance of power to R_L, i.e.:

$$N_{line} = kT_{ph}B - kT_{ph}BG_l \tag{12.41}$$

Dividing equation (12.41) by kB gives:

$$T_{line} = T_{ph}(1 - G_l) \tag{12.42}$$

Referring the line temperature to the source terminals then gives equation (12.37).

EXAMPLE 12.3

Figure 12.15 shows a simple superheterodyne receiver consisting of a front end low noise amplifier (LNA), a mixer and two stages of IF amplification. The source temperature of the receiver is 100 K and the characteristics of the individual receiver subsystems are:

Device	*Gain (or conversion loss)*	T_e
LNA	12 dB	50 K
Mixer	–6 dB	
IF Amp 1	20 dB	1000 K
IF Amp 2	30 dB	1000 K

Calculate the noise power at the output of the second IF amplifier in a 5.0 MHz bandwidth.

Using equation (12.34):

$$T_{syst\ in} = T_s + T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1G_2} + \frac{T_{e4}}{G_1G_2G_3}$$

where:

$G_1 = 12$ dB $= 15.8$, $G_2 = -6$ dB $= 0.25$ $(= G_l)$, $G_3 = 20$ dB $= 100$

$T_s = 100$ K, $T_{e1} = 50$ K, $T_{e2} = T_{ph}(1 - G_l)/G_l = 290(1 - 0.25)/0.25 = 870$ K, $T_{e3} = 1000$ K, $T_{e4} = 1000$ K, (Notice that the physical temperature of the mixer has been assumed to be 290 K. This is the normal assumption unless information to the contrary is given. Notice also that the gain of the final stage is irrelevant as far as the system equivalent input temperature is concerned.)

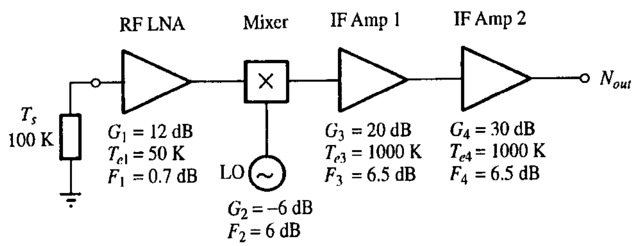

Figure 12.15 *Superheterodyne receiver.*

Therefore:

$$T_{syst\ in} = 100 + 50 + \frac{870}{15.8} + \frac{1000}{15.8 \times 0.25} + \frac{1000}{15.8 \times 0.25 \times 100}$$

$$= 100 + 50 + 55 + 253 + 3 = 461 \text{ K}$$

If the effective noise bandwidth of the system (often determined by the final IF amplifier) is 5.0 MHz then the total noise output power will be:

$$N_{out} = kT_{syst\ in}BG_1G_2G_3G_4$$

$$= 1.38 \times 10^{-23} \times 461 \times 5 \times 10^6 \times 15.8 \times 0.25 \times 100 \times 1000$$

$$= 1.26 \times 10^{-8} \text{ W} = -49.0 \text{ dBm}$$

It is interesting to repeat this calculation with the positions of the LNA and mixer reversed. In this case:

$$T_{syst\ in} = 100 + 870 + \frac{50}{0.25} + \frac{1000}{0.25 \times 15.8} + \frac{1000}{0.25 \times 15.8 \times 100}$$

$$= 100 + 870 + 200 + 253 + 3 = 1426 \text{ K}$$

The equivalent system input noise temperature, and therefore the total system output noise, has been degraded by a factor of 3. The reason for the presence of a low noise amplifier at the front end of a receiver thus becomes obvious.

12.3.3 Noise factor and noise figure

The noise factor, f, of an amplifier is defined as the ratio of SNR at the system input to SNR at the system output *when the input noise corresponds to a temperature of 290 K*, i.e.:

$$f = \left. \frac{(S/N)_i}{(S/N)_o} \right|_{N_i = k\,290\,B} \tag{12.43}$$

where:

$$(S/N)_i = \frac{\text{signal power at input, } S_i}{k\ 290\ B} \tag{12.44(a)}$$

and:

$$(S/N)_o = \frac{\text{signal power at output, } S_o}{k(290 + T_e)B\ G} \tag{12.44(b)}$$

The specification of the input noise temperature as 290 K allows all devices and systems to be compared fairly, on the basis of their quoted noise factor. f is therefore a figure of merit for comparing the noise performance of different devices and systems. Substituting equations (12.44) into (12.43):

$$f = \frac{S_i\ /\ (k\ 290\ B)}{GS_i\ /\ [G\ k(290 + T_e)B]} = \frac{290 + T_e}{290} \tag{12.45(a)}$$

i.e.:

$$f = 1 + \frac{T_e}{290} \tag{12.45(b)}$$

It is important to remember that, strictly speaking, f is only the ratio of input to output SNR if:

1. The device is operating at its equilibrium temperature in an ambient (290 K) environment. (This is necessary for the quoted T_e to be reliable.)
2. The source temperature at the device input is 290 K.

In practice it is the second condition which is most likely to be unfulfilled. Despite the arbitrary nature of the assumed source temperature in the definition of f, it is still possible to make accurate calculations of overall system noise temperature (and therefore overall SNRs) even when $T_s \neq 290$ K. This is because the noise factor of any device can be converted to an equivalent noise temperature using:

$$T_e = (f - 1)290 \quad \text{(K)} \tag{12.46}$$

and equivalent noise temperature makes no assumption at all about source temperature. If preferred, however, the noise effects due to several subsystems can be cascaded before converting to noise temperatures. Comparing equations (12.34) and (12.35) the equivalent noise temperature of the cascaded subsystems is:

$$T_e = T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1 G_2} \quad \text{(K)} \tag{12.47}$$

This can be rewritten in terms of noise factors as:

$$(f - 1)290 = (f_1 - 1)290 + \frac{(f_2 - 1)290}{G_1} + \frac{(f_3 - 1)290}{G_1 G_2} \tag{12.48}$$

Dividing by 290 and adding 1:

$$f = f_1 + \frac{f_2 - 1}{G_1} + \frac{f_3 - 1}{G_1 G_2} \tag{12.49}$$

This is called the Friis noise formula. The final system noise temperature (referred to the input of device 1) is then calculated from equations (12.46) and (12.35).

Traditionally noise factor is quoted in decibels, i.e.:

$$F = 10 \log_{10} f \quad \text{(dB)} \tag{12.50}$$

and in this form is called the *noise figure*. It is therefore essential to remember to convert F to a ratio before using it in the calculations described above. (Care is required since the terms noise figure and noise factor and the symbols f and F are often used interchangeably in practice.) Table 12.2 gives some noise figures, noise factors and their equivalent noise temperatures.

The noise figure of lossy devices (such as transmission lines or passive mixers) is related to their 'gain', G_l, by:

$$f = 1/G_l \tag{12.51}$$

where f and G_l are expressed as ratios, or alternatively by:

$$F = -G_l \quad \text{(dB)} \tag{12.52}$$

where F and G_l are expressed in decibels. A transmission line with 10% power loss (i.e. $G_l = 0.9$) therefore has a noise factor given by:

$$f = 1/G_l = 1/0.9 = 1.11 \tag{12.53}$$

(or 0.46 dB as a noise figure). A mixer with a conversion loss of 6 dB (i.e. a conversion gain of –6 dB) has a noise figure given by:

$$F = -G_l = 6 \text{ dB} \tag{12.54}$$

(or 4.0 as a noise factor). Strictly speaking a mixer would have slightly greater noise figure than this due to the contribution of non-thermal noise by the diodes in the mixer circuit. For passive mixers such non-thermal effects can usually be neglected. For active mixers this might not be so (although in this case there would probably be a conversion gain rather than loss).

Table 12.2 *Comparison of noise performance measures.*

T_e	f	F	*Comments*
0 K	1.00	0 dB	Perfect (i.e. noiseless) device
10 K	1.03	0.2 dB	Excellent LNA
100 K	1.34	1.3 dB	Good LNA
290 K	2.00	3.0 dB	Typical LNA
500 K	2.72	4.4 dB	Typical amplifier
1000 K	4.45	6.5 dB	Poor quality amplifier
10,000 K	35.50	15.5 dB	Temperature of a noise source

EXAMPLE 12.4

The output noise of the system shown in Figure 12.15 is now recalculated using noise factors instead of noise temperatures.

The noise factor of the entire system is given in equation (12.49) by:

$$\begin{aligned} f &= f_1 + \frac{(f_2-1)}{G_1} + \frac{(f_3-1)}{G_1G_2} + \frac{(f_4-1)}{G_1G_2G_3} \\ &= 10^{0.7/10} + \frac{10^{6/10}-1}{15.8} + \frac{10^{6.5/10}-1}{15.8\times 0.25} + \frac{10^{6.5/10}-1}{15.8\times 0.25\times 100} \\ &= 1.17 + \frac{2.98}{15.8} + \frac{3.47}{15.8\times 0.25} + \frac{3.47}{15.8\times 0.25\times 100} \\ &= 1.17 + 0.19 + 0.88 + 0.01 = 2.25 \end{aligned}$$

(or $F = 3.5$ dB)

The equivalent system noise temperature at the input to the low noise amplifier (LNA) is:

$$T_e = (f - 1)290$$
$$= (2.25 - 1)290 = 362 \text{ K}$$

and the total noise power at the system output in a bandwidth of 5.0 MHz is therefore:

$$N = k(T_s + T_e)BG_{syst}$$
$$= 1.38 \times 10^{-23} \times (100 + 362) \times 5 \times 10^6 \times 10^{56/10}$$
$$= 1.27 \times 10^{-8} \text{ W} = -49.0 \text{ dBm}$$

12.4 Radio communication link budgets

A communication system link budget refers to the calculation of received signal-to-noise ratio given a specification of transmitted power, transmission medium attenuation and/or gain, and all sources of noise.

The calculation is often set out systematically (in a similar way to a financial budget) accounting explicitly for the various sources of gain, attenuation and noise. For single section line communications the essential elements are transmitted power, cable attenuation, receiver gain and noise figure. The calculation is then simply a matter of applying the Friis formula, or its equivalent, as described in sections 12.3 – 12.3.3. For radio systems the situation is different in that signal energy is lost not only as a result of attenuation (i.e. energy which is dissipated as heat) but also due to its being radiated in directions other than directly towards the receiving antenna. For multi-section communication links the effects of analogue, amplifying, repeaters can be accounted for using their gains and noise figures in the Friis formula, and the effects of digital regenerative repeaters can be accounted for by summing the BERs of each section (providing the BER is small, see section 6.3). Before the details of a radio communications link budget are described some important antenna concepts are reviewed.

12.4.1 Antenna gain, effective area and efficiency

An isotropic antenna (i.e. one which radiates electromagnetic energy equally well in all directions), radiating P_{rad} W of power, supports a power density at a distance R (Figure 12.16) given by:

$$W_{isotrope} = \frac{P_{rad}}{4\pi R^2} \quad (\text{W m}^{-2}) \tag{12.55}$$

The observation point is assumed, here, to be in the far-field of the antenna, i.e. $R \geq 2D^2/\lambda$ where D is the largest dimension (often the diameter) of the antenna. In this region the reactive and radiating antenna near-fields are negligible, and the radiating far-

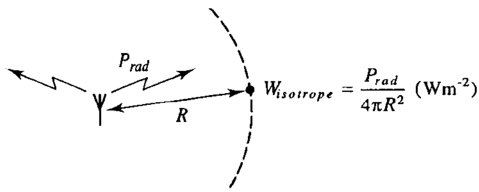

Figure 12.16 *Power density radiated by an isotropic antenna.*

field can be assumed to be a transverse electromagnetic (TEM) wave. The radiating far-field pattern is independent of distance and, in any small region of space, the wavefront can be considered approximately plane. The field strength in the antenna far-field is related to power density by:

$$\frac{E_{RMS}^2}{Z_o} = W \quad (\text{W/m}^2) \tag{12.56}$$

where $Z_o = E/H = 377\ \Omega$ is the plane wave impedance of free space. Since $1/(4\pi R^2)$ in equation (12.55) represents a purely geometrical dilution of power density as the spherical wave expands, this factor is often referred to as the *spreading loss.* The *radiation intensity*, I, in an isotropic antenna's far-field is given by:

$$I_{isotrope} = \frac{P_{rad}}{4\pi} \quad (\text{W steradian}^{-1} \text{ or W rad}^{-2}) \tag{12.57}$$

The mutually orthogonal requirement on $\mathbf{E}, \mathbf{H}$ and $\mathbf{k}$ in the far-field (where $\mathbf{E}$ is electric field strength, $\mathbf{H}$ is magnetic field strength and $\mathbf{k}$ is vector wave number pointing in the direction of propagation) excludes the possibility of realising isotropic radiators. Thus all practical antennas radiate preferentially in some directions over others. If radiation intensity $I(\theta, \phi)$ is plotted against spherical coordinates θ and ϕ the resulting surface (i.e. radiation pattern) will be spherical for a (hypothetical) isotrope and non-spherical for any realisable antenna, Figure 12.17. The gain, $G(\theta, \phi)$, of a transmitting antenna can be defined in azmuth θ and elevation ϕ as the ratio of radiation intensity in the direction θ, ϕ to the radiation intensity observed by replacing the antenna with a *lossless* isotrope:

$$G(\theta, \phi) = \frac{I(\theta, \phi)}{I_{lossless\ isotrope}} = \frac{I(\theta, \phi)}{P_T/(4\pi)} \tag{12.58(a)}$$

where P_T is the transmitter output power and the antenna is assumed to be well matched to its transmission line feed. A related quantity, antenna directivity, $D(\theta, \phi)$, can be defined by:

$$D(\theta, \phi) = \frac{I(\theta, \phi)}{I_{isotrope}}$$

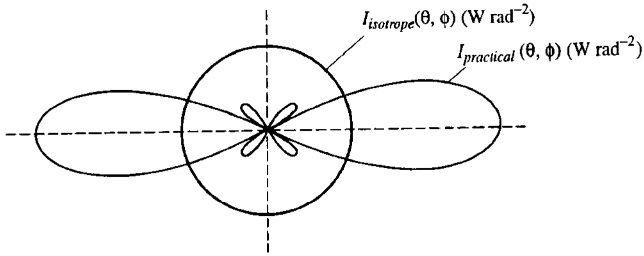

Figure 12.17 *Two-dimensional polar plots of antenna radiation intensity for isotropic and practical antenna.*

$$= \frac{I(\theta, \phi)}{P_{rad}/(4\pi)} \tag{12.58(b)}$$

where the hypothetical isotropic antenna is now assumed to have a loss equal to that of the actual antenna. Directivity and gain therefore have identical shape and are related by:

$$G(\theta, \phi) = \eta_\Omega D(\theta, \phi) \tag{12.58(c)}$$

where $\eta_\Omega = P_{rad}/P_T$ is called the ohmic efficiency of the antenna. The ohmic losses relate physically to the electromagnetic energy dissipated by induced conduction currents flowing in the metallic conductors of the antenna and induced displacement currents flowing in the dielectric of the antenna. (The former usually dominate the latter and are often referred to as I^2R losses.) As an alternative both antenna gain and directivity can be defined as ratios of far-field power densities, i.e.:

$$G(\theta, \phi) = \frac{W(\theta, \phi, R)}{W_{lossless\ isotrope}(R)} \tag{12.59(a)}$$

and

$$D(\theta, \phi) = \frac{W(\theta, \phi, R)}{W_{isotrope}(R)} \tag{12.59(b)}$$

Using equations (12.55), (12.58), (12.59) and $P_{rad} = \eta_\Omega P_T$, the power density at a distance R from a transmitting antenna can be found from either of the following formulas:

$$W(\theta, \phi, R) = \frac{P_{rad}}{4\pi R^2} D(\theta, \phi) \quad (\text{W/m}^2) \tag{12.60(a)}$$

$$W(\theta, \phi, R) = \frac{P_T}{4\pi R^2} G(\theta, \phi) \quad (\text{W/m}^2) \tag{12.60(b)}$$

Figure 12.18 shows a typical, Cartesian coordinate, directivity pattern for a microwave reflector antenna. The effective area, a_e, of a receiving antenna, is defined as the ratio of carrier power C received at the antenna terminals, when the antenna is illuminated by a plane wave from the direction θ, ϕ, to the power density in the plane wave (Figure 12.19),

i.e.:

$$a_e(\theta, \phi) = \frac{C(\theta, \phi)}{W} \quad (\text{m}^2) \tag{12.61}$$

(The plane wave is assumed to be polarisation matched to the antenna.) Unsurprisingly, there is an intimate connection between an antenna's gain and its effective area which can be expressed by:

$$G(\theta, \phi) = \frac{4\pi a_e(\theta, \phi)}{\lambda^2} \tag{12.62}$$

Equation (12.62) is a widely used expression of antenna reciprocity which can be derived from the Lorentz reciprocity theorem [Collin]. Usually interest is focused on the gain and effective area corresponding to the direction of maximum radiation intensity. This direction is called antenna boresight and is always implied if θ and ϕ are not specified.

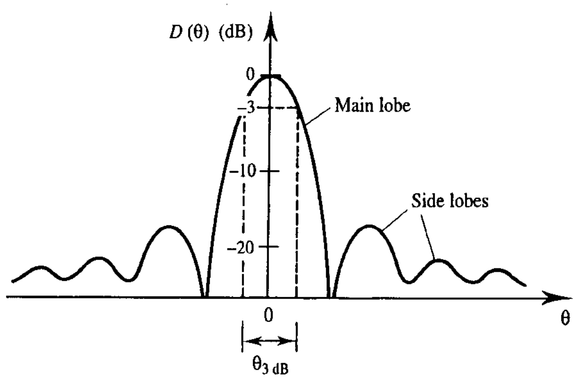

Figure 12.18 *Two-dimensional Cartesian plot of directivity (in dB) for microwave antenna (3 dB beamwidth of axisymmetric reflectors may be estimated using θ_{3dB} = 1.2 λ/D rad).*

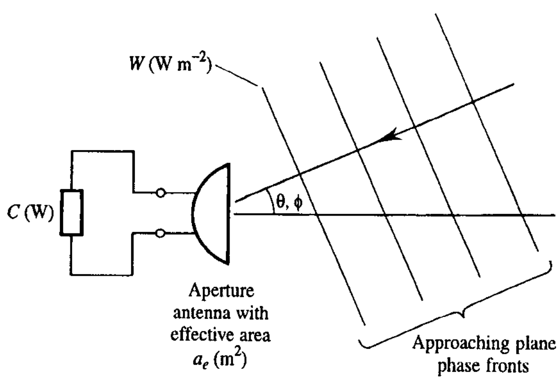

Figure 12.19 *Aperture antenna receiving plane wave from θ, ϕ direction.*

For aperture antennas, such as paraboloidal reflectors, it is intuitively reasonable to associate a_e, at least approximately, with the physical antenna aperture. Unfortunately, they cannot be assumed to be identical since not all parts of the antenna aperture are fully utilised. There are several reasons for this including non-uniform illumination of the aperture by the feed, non-zero illumination of the region outside the aperture by the feed, aperture blockage caused by the presence of the feed and feed struts, and random errors in the definition of the antenna's reflecting surface. These and other effects can be accounted for by reducing the physical area, A_{ph}, of an aperture antenna by an aperture efficiency factor, η_{ap}. This gives the antenna's *effective* aperture, A_e, i.e.:

$$A_e = \eta_{ap} A_{ph} \quad (\text{m}^2) \tag{12.63(a)}$$

These aperture losses are in addition to the ohmic losses referred to earlier. Ohmic loss is accounted for by the ohmic efficiency, which when applied to the effective aperture gives the antenna's effective area as defined in equation (12.61), i.e.:

$$a_e = \eta_\Omega A_e \quad (\text{m}^2) \tag{12.63(b)}$$

The effective area of the antenna is therefore related to its physical area by:

$$a_e = \eta_{ap} \eta_\Omega A_{ph} \quad (\text{m}^2) \tag{12.64}$$

η_{ap} and η_Ω can be further broken down into more specific sources of loss. Any more detailed accounting for losses, however, is usually important only for antenna designers and is of rather academic interest to communications systems engineers. (Care should be taken in interpreting antenna efficiencies, however, since the terms effective aperture and effective area are often used interchangeably, as are the terms aperture efficiency and antenna efficiency.)

It is important to realise that antenna aperture efficiency is certain to be less than unity only for antennas with a well defined aperture of significant size (in terms of wavelengths). Wire antennas, such as dipoles, for example, have aperture efficiencies greater than unity if A_{ph} is taken to be the area presented by the wire to the incident wavefront.

EXAMPLE 12.5

The power density, radiation intensity and electric field strength are now calculated at a distance of 20 km from a microwave antenna having a directivity of 42.0 dB, an ohmic efficiency of 95% and a well matched 4 GHz transmitter with 25 dBm of output power.

The gain is given by:

$$\begin{aligned} G &= \eta_\Omega D \\ &= 10 \log_{10} \eta_\Omega + D_{dB} \quad (\text{dB}) \\ &= 10 \log_{10} 0.95 + 42.0 = -0.2 + 42.0 = 41.8 \text{ dB} \end{aligned}$$

and, from equation (12.60(b)), the received power density is given by:

$$
\begin{aligned}
W &= \frac{P_T}{4\pi R^2} G_T \\
&= P_T + G_T - 10 \log_{10} (4\pi R^2) \ \text{dBm/m}^2 \\
&= 25 + 41.8 - 10 \log_{10} (4\pi \times 20,000^2) \\
&= -30.2 \ \text{dBm/m}^2 = -60.2 \ \text{dBW/m}^2 \\
&= 9.52 \times 10^{-7} \ \text{W/m}^2
\end{aligned}
$$

The radiation intensity is given by equation (12.58(b)) as:

$$
\begin{aligned}
I &= \frac{P_{rad}}{4\pi} D = \frac{\eta_\Omega P_T}{4\pi} D \\
&= \frac{0.95 \times 10^{\frac{25}{10}}}{4\pi} \times 10^{\frac{42}{10}} \ \text{mW/rad}^2 \\
&= 3.789 \times 10^5 \ \text{mW/rad}^2
\end{aligned}
$$

and from equation (12.56):

$$
E_{RMS} = \sqrt{W Z_o} = \sqrt{9.52 \times 10^{-7} \times 377} \ \text{V/m} = 18.9 \ \text{mV/m}
$$

If an identical receiving antenna is located 20 km from the first, the available carrier power, C, at its terminals could be calculated as follows:

$$
\lambda = \frac{c}{f} = \frac{3 \times 10^8}{4 \times 10^9} = 0.075 \ \text{m}
$$

The antenna effective area, equation (12.62), is:

$$
\begin{aligned}
a_e &= G \frac{\lambda^2}{4\pi} = 10^{\frac{41.8}{10}} \left[\frac{(0.075)^2}{4\pi} \right] \\
&= 6.775 \ \text{m}^2
\end{aligned}
$$

and from equation (12.61):

$$
\begin{aligned}
C &= W a_e \\
&= 9.52 \times 10^{-7} \times 6.775 = 6.45 \times 10^{-6} \ \text{W} = -21.9 \ \text{dBm}
\end{aligned}
$$

12.4.2 Free space and plane earth signal budgets

Consider the free space radio communication link shown in Figure 12.20. The power density at the receiver radiated by a lossless isotrope would be:

$$
W_{lossless\ isotrope} = \frac{P_T}{4\pi R^2} \ (\text{W m}^{-2}) \tag{12.65}
$$

For a practical transmitting antenna with gain G_T the power density at the receiver is actually:

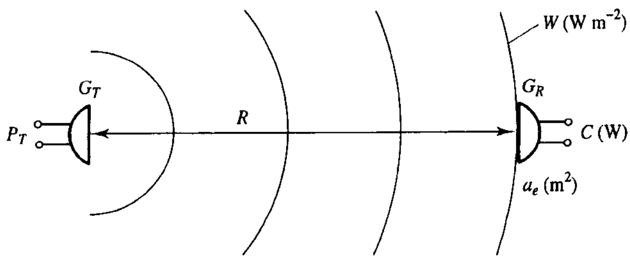

Figure 12.20 *Free space propagation.*

$$W = \frac{P_T}{4\pi R^2} G_T \quad (\mathrm{W\ m^{-2}}) \tag{12.66}$$

The carrier power available at the receive antenna terminals is given by:

$$C = W\, a_e \quad (\mathrm{W}) \tag{12.67}$$

which, on substituting equation (12.66) for W, gives:

$$C = \frac{P_T}{4\pi R^2} G_T\, a_e \quad (\mathrm{W}) \tag{12.68}$$

Using antenna reciprocity (equation (12.62)):

$$C = \frac{P_T}{4\pi R^2} G_T G_R \frac{\lambda^2}{4\pi} \quad (\mathrm{W}) \tag{12.69}$$

where G_R is the gain of the receiving antenna. (Note that $\lambda^2/4\pi$ can be identified as the effective area, a_e, of a lossless isotrope.) Equation (12.69) can be rewritten as:

$$C = P_T G_T \left(\frac{\lambda}{4\pi R} \right)^2 G_R \quad (\mathrm{W}) \tag{12.70}$$

This is the basic free space transmission loss formula for radio systems. The quantity $P_T G_T$ is called the effective isotropic radiated power (EIRP) and the quantity $[\lambda/(4\pi R)]^2$ is called the free space path loss (FSPL). Notice that FSPL is a function of wavelength. This is because it contains a factor to convert the receiving antenna effective area to gain in addition to the (geometrical) spreading loss.

Equation (12.70) is traditionally expressed using decibel quantities:

$$C = \mathrm{EIRP} - \mathrm{FSPL} + G_R \quad (\mathrm{dBW}) \tag{12.71(a)}$$

where:

$$\mathrm{EIRP} = 10 \log_{10} P_T + 10 \log_{10} G_T \quad (\mathrm{dBW}) \tag{12.71(b)}$$

and:

$$\mathrm{FSPL} = 20 \log_{10} \left(\frac{4\pi R}{\lambda} \right) \quad (\mathrm{dB}) \tag{12.71(c)}$$

(The same symbol is used here for powers and gains whether they are measured in natural units or decibels. The context, however, leaves no doubt as to which is intended.)

Figure 12.21 shows a radio link operating above a plane earth. In this case there are two possible propagation paths between the transmitter and receiver, one direct and the other reflected via the ground. Assuming a complex (voltage) reflection coefficient at the ground, $\rho e^{j\phi}$, then the field strength, E, at the receiver will be changed by a (complex) factor, F, given by:

$$F = 1 + \rho e^{j\phi} e^{-j2\pi(d_2 - d_1)/\lambda} \tag{12.72}$$

where d_1 and d_2 are the lengths of the direct and reflected paths respectively and $2\pi(d_2 - d_1)/\lambda = \theta$ is the resulting excess phase shift of the reflected, over the direct, path. Figure 12.22 illustrates the (normalised) phasor addition of the direct and reflected fields. Assuming perfect reflection at the ground (i.e. $\rho e^{j\phi} = -1$) and using $e^{-j\theta} = \cos\theta - j\sin\theta$ the magnitude of the field strength gain factor can be written, for practical multipath geometries (see Problem 12.8), as:

$$|F| = 2\sin\left(\frac{2\pi h_T h_R}{\lambda R}\right) \tag{12.73}$$

where h_T and h_R are transmit and receive antenna heights respectively and R is the horizontal distance between transmitter and receiver, Figure 12.21. The power density at the receiving antenna aperture, and therefore the received power, is increased over that for free space by a factor $|F|^2$. Thus, using natural (not decibel) quantities:

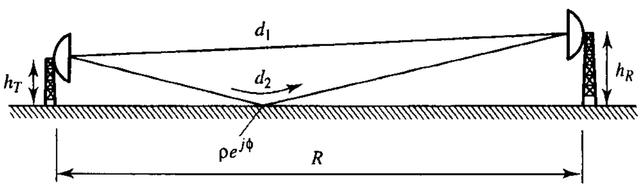

Figure 12.21 *Propagation over a plane earth.*

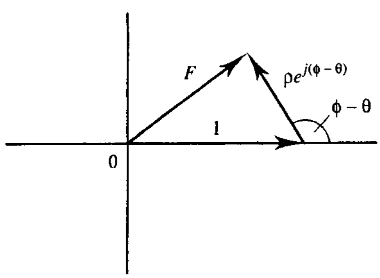

Figure 12.22 *Phasor addition of (normalised) fields from direct and reflected paths in Figure 12.21.*

$$C = \text{EIRP} \times \text{FSPL} \times |F|^2 \times G_R$$
$$= P_T G_T \left(\frac{\lambda}{4\pi R} \right)^2 4 \sin^2 \left(\frac{2\pi h_T h_R}{\lambda R} \right) G_R \quad \text{(W)} \tag{12.74}$$

A typical interference pattern resulting from the two paths is illustrated as a function of horizontal range and height in Figure 12.23(a) and Figure 12.23(b) respectively. Equation (12.74) can be expressed in decibels as:

$$C = P_T + G_T - \text{FSPL} + G_R + 6.0 + 20 \log_{10} \left| \sin \left(\frac{2\pi h_T h_R}{\lambda R} \right) \right| \quad \text{(dBW)} \tag{12.75}$$

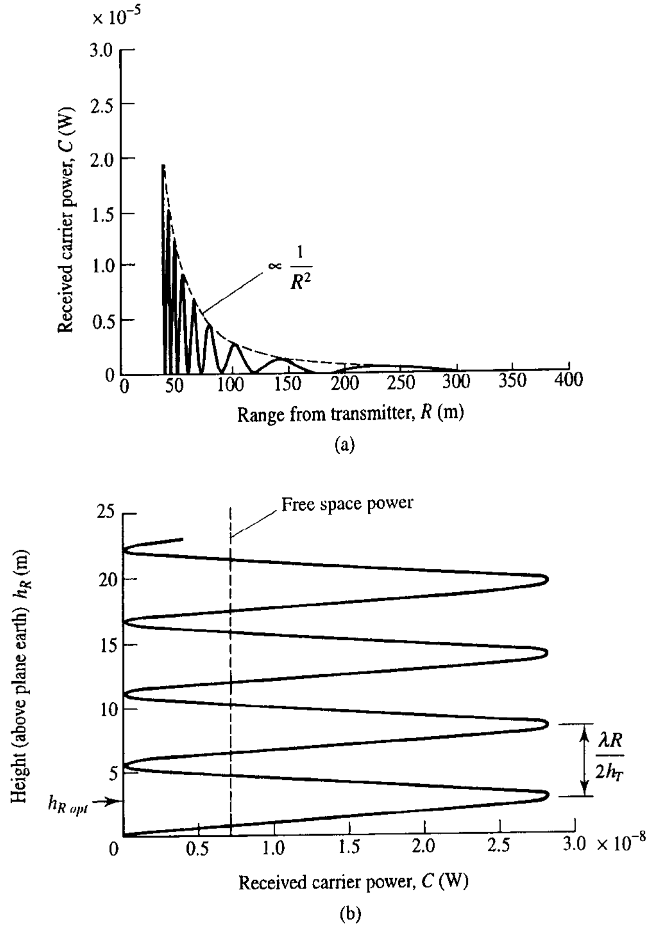

Figure 12.23 *Variation of received power with: (a) distance at a height of 2 m; (b) height at a range of 1000 m from transmitter. (Parameters as used in equation (12.74): $P_T G_T = 10$ W, $\lambda = 0.333$ m, $h_T = 30$ m, $G_R = 0$ dB.)*

Figure 12.24 shows the variation of received power on a dB scale with respect to transmitter–receiver distance. (A plot of power density in dBW m^{-2} or electric field in dBμV m^{-1} has identical shape.) Notice that the peaks, corresponding to points of constructive interference, are 6 dB above the free space value and the troughs, corresponding to points of destructive interference, are, in principle, ∞ dB below the free space value. (In practice the minima do not represent zero power density or field strength because $\rho < 1.0$.) The furthest point of constructive interference occurs for the smallest argument of sin $[(2\pi h_T h_R)/(\lambda R)]$ which gives 1.0, i.e. when:

$$\frac{2\pi h_T h_R}{\lambda R} = \frac{\pi}{2} \qquad (12.76(a))$$

The distance R_{max}, to which this corresponds, is:

$$R_{max} = \frac{4 h_T h_R}{\lambda} \quad (m) \qquad (12.76(b))$$

For $R > R_{max}$ the total field decays monotonically. When $2\pi h_T h_R / \lambda R$ is small the approximation $\sin x \approx x$ can be used and the received power becomes:

$$C = P_T G_T \left(\frac{h_T h_R}{R^2} \right)^2 G_R \quad (W) \qquad (12.77)$$

or, in decibels:

$$C = P_T + G_T - \text{PEPL} + G_R \quad (\text{dBW}) \qquad (12.78(a))$$

where PEPL, the plane earth path loss, is given by:

$$\text{PEPL} = 20 \log_{10} \left(\frac{R^2}{h_T h_R} \right) \quad (\text{dB}) \qquad (12.78(b))$$

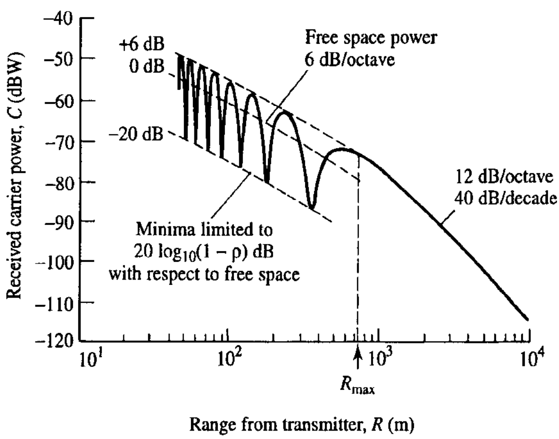

Figure 12.24 *Variation of received power with range for plane earth propagation. (Parameters used are the same as for Figure 12.23.)*

Equation (12.77) shows that over a plane conducting earth the received power, and therefore the power density at the receiving antenna aperture, decays as $1/R^4$. The corresponding field strength therefore decays as $1/R^2$.

The optimum receive antenna height, $h_{R\,opt}$, for fixed λ, R and h_T is found by differentiating $|F|$ in equation (12.73) with respect to h_R and setting the result to zero, i.e.:

$$\frac{d|F|}{dh_R} = 2\cos\left(\frac{2\pi h_T h_R}{\lambda R}\right)\frac{2\pi h_T}{\lambda R} = 0 \tag{12.79}$$

This requires that:

$$\frac{2\pi h_T h_{R\,opt}}{\lambda R} = \frac{n\pi}{2}, \quad n = 1, 3, 5, \cdots \tag{12.80(a)}$$

i.e.:

$$h_{R\,opt} = n\frac{\lambda R}{4h_T} \quad \text{(m)} \tag{12.80(b)}$$

where n is an odd integer. (A second derivative $(d^2|F|/dh_R^2)$ confirms that this is a condition for maximum, rather than minimum, received power.) Clearly antenna heights would not normally be chosen to be higher than necessary and so, in the absence of other considerations, n would be chosen to be equal to 1. (It would actually be useful to optimise an antenna height only if the ground reflection was reasonably stable in both magnitude *and phase* – a condition not often encountered.)

EXAMPLE 12.6

A 6 GHz, 40 km, LOS link uses 2.0 m axisymmetric paraboloidal reflectors for both transmitting and receiving antennas. The ohmic and aperture efficiencies of the antennas are 99% and 70% respectively and both antennas are mounted at a height of 25 m. The transmitter power is 0 dBW. Find the received power for both free space and plane earth conditions. Which condition is most likely to prevail if the path is over water?

$$\lambda = \frac{c}{f} = \frac{3\times10^8}{6\times10^9} = 0.05 \text{ m}$$

From equation (12.64) we have:

$$\begin{aligned} a_e &= \eta_\Omega \eta_{ap} A_{ph} \\ &= 0.99\times0.70\times\pi\times1.0^2 = 2.177 \text{ m}^2 \end{aligned}$$

and using equation (12.62):

$$G = \frac{4\pi a_e}{\lambda^2} = \frac{4\pi\times2.177}{0.05^2} = 1.094\times10^4 = 40.4 \text{ dB}$$

$$\text{FSPL} = 20\log_{10}\left(\frac{4\pi R}{\lambda}\right)$$

$$= 20\log_{10}\left(\frac{4\pi 40\times 10^3}{0.05}\right) = 140.0 \text{ dB}$$

For free space conditions equation (12.71(a)) gives:

$$C = P_T + G_T - \text{FSPL} + G_R$$

$$= 0 + 40.4 - 140.0 + 40.4 = -59.2 \text{ dBW}$$

For plane earth conditions (assuming perfect reflection) equation (12.76(b)) gives:

$$R_{max} = \frac{4h_T h_R}{\lambda}$$

$$= \frac{4\times 25^2}{0.05} = 50{,}000 \text{ m}$$

The receive antenna is thus closer to the transmit antenna than the furthest point of constructive interference and equation (12.75) rather than (12.78) must therefore be used.

$$C = C\,|_{free\ space} + 6.0 + 20\log_{10}\left|\sin\left(\frac{2\pi h_T h_R}{\lambda R}\right)\right|$$

$$= -59.2 + 6.0 + 20\log_{10}\left|\sin\left(\frac{2\pi\times 25^2}{0.05\times 40\times 10^3}\right)\right|$$

$$= -59.2 + 6.0 - 0.7 = -53.9 \text{ dBW}$$

Note that this is only 0.7 dB below the maximum possible received power so the antenna heights must be close to optimum for plane earth propagation. Equation (12.80) shows that optimum (equal) antenna heights would be given by:

$$h_R h_T = \frac{\lambda R}{4} = \frac{0.05\times 40\times 10^3}{4} = 500 \text{ m}^2$$

i.e.:

$$h_R = h_T = \sqrt{500} = 22.36 \text{ m}$$

As has been said in the text, however, it is unlikely that such precise positioning of antennas would be of benefit in practice. To assess whether free space or plane earth propagation is likely to occur, a first order calculation based on antenna beamwidth can be carried out as follows:

$$\text{Antenna beamwidth} \approx 1.2\,\frac{\lambda}{D} \text{ rad}$$

$$= 1.2\left(\frac{0.05}{2.0}\right) = 0.03 \text{ rad } (= 1.7 \text{ degrees})$$

Under normal atmospheric conditions for a horizontal path over water, with equal height transmit and receive antennas, the point of specular reflection will be half way along the link. The vertical width of the antenna beam (between –3 dB points) is given by:

$$\text{Vertical width} \approx \text{beamwidth} \times \tfrac{1}{2} \text{ path length}$$

$$= 0.03\times \tfrac{1}{2}\times 40\times 10^3 = 600 \text{ m}$$

Since half this is much greater than the path clearance and since water is a good reflector at microwave frequencies then specular reflection (for a calm water surface) is likely to be strong and a plane earth calculation is appropriate.

12.4.3 Antenna temperature and radio noise budgets

The overall noise power in a radio communications receiver depends not only on internally generated receiver noise but also on the electromagnetic noise collected by the receiver's antenna. Just as the equivalent thermal noise of a circuit or receiver subsystem can be represented by a noise temperature, so too can the noise received by an antenna. The equivalent noise temperature of a receiving antenna, T_{ant}, is defined by:

$$T_{ant} = \frac{\text{available NPSD at antenna terminals}}{\text{Boltzmann's constant}} = \frac{G_N(f)}{k} \quad \text{(K)} \tag{12.81}$$

where $G_N(f)$ is assumed to be white and the noise power spectral density is one sided. Antenna noise originates from several different sources. Below about 30 MHz it is dominated by the broadband radiation produced in lightning discharges associated with thunderstorms. This radiation is trapped by the ionosphere and so propagates world-wide. Such noise is sometimes referred to as atmospherics. The ionosphere is essentially transparent above about 30 MHz and between this frequency and 1 GHz the dominant noise is galactic. This has a steeply falling spectral density with increasing frequency (the slope is about –25 dBK/decade). It arises principally due to synchrotron radiation produced by fast electrons moving through the galactic magnetic field. Because the galaxy is very oblate in shape, and also because the earth is not located at its centre, galactic noise is markedly anisotropic and is much greater when the receiving antenna is pointed towards the galactic centre than when it is pointed to the galactic pole.

Above 1 GHz galactic noise is relatively weak. This leaves atmospheric thermal radiation and ground noise as the dominant noise processes. Atmospheric and ground noise is approximately flat with frequency up to about 10 GHz, its spectral density depending sensitively on antenna elevation angle. As elevation increases from 0° to 90° the thickness of atmosphere through which the antenna beam passes decreases as does the influence of the ground, both effects leading to a decrease in received thermal noise. In this frequency range a zenith-pointed antenna during clear sky conditions may have a noise temperature close to the cosmic background temperature of 3 K. Above 10 GHz resonance effects (of water vapour molecules at 22 GHz and oxygen molecules at 60 GHz) lead to increasing atmospheric attenuation and, therefore, thermal noise emission. The typical 'clear sky' noise temperature, as would be measured by a lossless narrow-beam antenna, is illustrated over a band of frequencies from HF to SHF in Figure 12.25.

In addition to noise received as electromagnetic radiation by the antenna, thermal noise will also be generated in the antenna itself. A simple equivalent circuit of an antenna is shown in Figure 12.26. The (total) antenna noise temperature, T_{ant}, is the sum

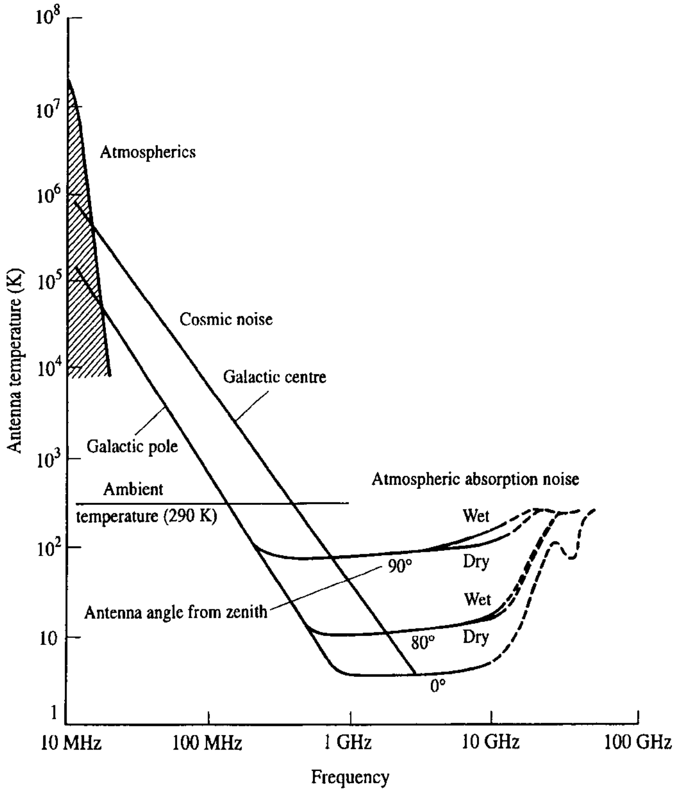

Figure 12.25 *Antenna sky noise temperature as a function of frequency and antenna elevation angle (source: Kraus, 1966, reproduced with his permission).*

of its aperture temperature, T_A (originating from external sources and reduced by ohmic losses) and its equivalent internal, thermal, temperature, i.e.:

$$T_{ant} = T_A \eta_\Omega + T_{ph}(1 - \eta_\Omega) \quad \text{(K)} \tag{12.82}$$

where η_Ω is the antenna ohmic efficiency and T_{ph} is the physical temperature of the antenna. For noise purposes, then, the radiation resistance, R_r, in Figure 12.26 can be assumed to be at the equivalent antenna aperture temperature and the ohmic resistance, R_Ω, can be assumed to be at the antenna physical temperature. The ohmic efficiency of the antenna is related to R_r and R_Ω by:

$$\eta_\Omega = \frac{R_r}{R_r + R_\Omega} \tag{12.83}$$

The calculation of T_A can be complicated but may be estimated using:

$$T_A = \frac{1}{4\pi} \int\int_{4\pi} D(\theta, \phi)\varepsilon(\theta, \phi)T_{ph}(\theta, \phi) \; d\Omega \quad \text{(K)} \tag{12.84}$$

where $\varepsilon(\theta, \phi)$ and $T_{ph}(\theta, \phi)$ are the emissivity and physical temperature, respectively, of the material lying in the direction θ, ϕ, and $d\Omega$ is an element of solid angle. (The emissivity of a surface is related to its voltage reflection coefficient, ρ, by $\varepsilon = 1 - \rho^2$ and the product $\varepsilon(\theta, \phi)T_{ph}(\theta, \phi)$ is sometimes called brightness temperature, $T_B(\theta, \phi)$.)

The quantity $D(\theta, \phi)\varepsilon(\theta, \phi)d\Omega$ in equation (12.84) can be interpreted as the fraction of power radiated by the antenna which is *absorbed* by the material lying in the direction between θ, ϕ and $\theta + d\theta$, $\phi + d\phi$. If the brightness temperature of the environment around the antenna changes discretely (assuming, for example, different, but individually uniform, temperatures for the ground, clear sky and sun) then equation (12.84) can be written in the simpler form:

$$T_A = \sum_i \alpha_i T_{ph,i} \qquad (12.85)$$

where α_i is the fraction of power radiated by the antenna which is absorbed by the ith body and $T_{ph,i}$ is the physical temperature of the ith body.

For more precise calculations still, scattering of radiation from one source by another (for example scattering of ground noise into the antenna by the atmosphere) must be accounted for. This level of detail, however, is usually the concern of radio remote sensing engineers rather than communications engineers.

Having established an antenna noise temperature the overall noise temperature of a radio receiving system is:

$$T_{syst\ in} = T_{ant} + T_e \qquad (12.86)$$

where T_e is the equivalent input noise temperature of the receiver.

12.4.4 Receiver equivalent input CNR

The equivalent input CNR of a radio receiver is given by combining equations (12.70) or (12.74) and (12.86) in the receiver bandwidth, i.e.:

$$\frac{C}{N} = \frac{C}{kT_{syst\ in}B} \qquad (12.87)$$

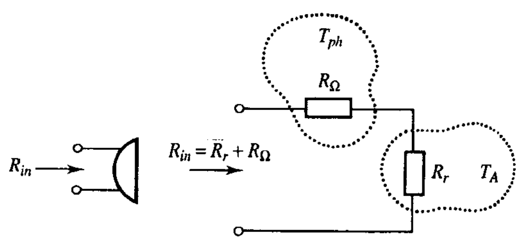

Figure 12.26 *Equivalent circuit of an antenna for noise calculations (R_r is radiation resistance, R_Ω is ohmic resistance and R_{in} is input resistance).*

The baseband BER and/or SNR of the demodulated signal are then found by applying any CNR detection gain provided by the demodulator and/or decoder, and finding the probability of bit error for the particular digital modulation scheme being used (including the mitigating effects of error control coding). If required the PCM decoded SNR of a baseband analogue signal can then be found by applying an equation such as (5.29).

It should be remembered that although the emphasis here has been on naturally occurring noise some systems are interference limited. Such interference may be random, quasi-periodic, intelligible (in which case it is usually called crosstalk) or a combination of these. Cellular radio is an example of a system which usually operates in an interference limited environment.

EXAMPLE 12.7

A 10 GHz terrestrial line of sight link has good clearance over rough terrain such that free space propagation can be assumed. The free space signal power at the receiving antenna terminals is –40.0 dBm. The overall noise figure of the receiver is 5.0 dB and the noise bandwidth is 20 MHz. Estimate the actual and effective clear sky CNRs at the antenna terminals assuming that the antenna has an ohmic efficiency of 95% and is at a physical temperature of 280 K. Also make a first order estimate of the effective CNR during a rain fade of 2 dB, assuming the rain is localised and occurs close to the receiving antenna.

From Figure 12.25 the clear sky aperture temperature, T_A, at 10 GHz for a horizontal link is about 100 K. The antenna temperature is given in equation (12.82) by:

$$T_{ant} = T_A \eta_\Omega + T_{ph}(1 - \eta_\Omega)$$

$$= 100 \times 0.95 + 280\,(1 - 0.95) = 95 + 14 = 109 \text{ K}$$

(Notice that the contribution from the physical temperature of the antenna is, in this case, probably within the uncertainty of the estimate of aperture temperature.) Now from equation (12.8):

$$N = kTB = 1.38 \times 10^{-23} \times 109 \times 20 \times 10^6 = 3.01 \times 10^{-14} \text{ W} = -135.2 \text{ dBW}$$

The actual clear sky CNR at the antenna terminals with a received power of –70 dBW is given by:

$$\frac{C}{N} = -70.0 - (-135.2) = 65.2 \text{ dB}$$

The equivalent noise temperature of the receiver is given by equation (12.46) as:

$$T_e = (f - 1)290 = (10^{\frac{5}{10}} - 1)290 = 627 \text{ K}$$

The system noise temperature (referred to the antenna output) is:

$$T_{syst\ in} = T_{ant} + T_e = 109 + 627 = 736 \text{ K}$$

The effective system noise power (referred to the antenna output) is:

$$N = kT_{syst\ in}B = 1.38 \times 10^{-23} \times 736 \times 20 \times 10^6$$

$$= 2.03 \times 10^{-13} \text{ W} = -126.9 \text{ dBW}$$

The effective carrier to noise ratio is:

$$\left.\frac{C}{N}\right|_{eff} = C - N \text{ (dB)}$$

$$= -70.0 - (-126.9) = 56.9 \text{ dB}$$

During a 2 dB rain fade carrier power will be reduced by 2 dB to −42 dBm and noise will be increased. A first order estimate of noise power during a fade can be found as follows.
Assuming the physical temperature of the rain (T_{rain}) is the same as that of the antenna then the aperture temperature may be recalculated using the transmission line of equation (12.42) as:

$$T_A = T_{sky} \times \text{fade} + T_{rain}(1 - \text{fade})$$

$$= 100 \times 10^{\frac{-2}{10}} + 280\,(1 - 10^{\frac{-2}{10}})$$

$$= 63 + 103 = 166 \text{ K}$$

$$T_{ant} = T_A\,\eta_\Omega + T_{ph}(1 - \eta_\Omega)$$

$$= 166 \times 0.95 + 280(1 - 0.95)$$

$$= 158 + 14 = 172 \text{ K}$$

$$N = k(T_{ant} + T_e)B$$

$$= 1.38 \times 10^{-23}\,(172 + 627) \times 20 \times 10^6$$

$$= 2.21 \times 10^{-13} \text{ W} = -126.6 \text{ dBW}$$

The effective CNR during the 2 dB fade is therefore:

$$\left.\frac{C}{N}\right|_{eff} = C - N$$

$$= -72.0 - (-126.6) = 54.6 \text{ dB}$$

The example shown above is intended to illustrate the concepts discussed in the text and probably contains spurious precision. The uncertainties associated with real systems mean that in practice a first order estimate of CNR would probably be based on a worst case antenna noise temperature of, say, 290 K. (The difference between this assumption and the above effective CNR calculation is only $10 \log_{10} ((290 + 627)/(172 + 627)) = 0.6$ dB.)

12.4.5 Multipath fading and diversity reception

Multipath fading occurs to varying extents in many different radio applications [Rummler, 1986]. It is caused whenever radio energy arrives at the receiver by more than one path. Figure 12.27 illustrates how multipath propagation may occur on a point-to-point line-of-sight microwave link. In this case multiple paths may occur due to ground reflections, reflections from stable tropospheric layers (with different refractive index) and refraction by tropospheric layers with extreme refractive index gradients. Other systems suffer multipath propagation due to the presence of scattering obstacles. This is the case for urban cellular radio systems for example.

There are two principal effects of multipath propagation on systems, their relative severity depending essentially on the relative bandwidth of the resulting channel compared with that of the signal being transmitted. For fixed point systems such as the microwave radio relay network the fading process is governed by changes in atmospheric conditions. Often, but not always, the spread of path delays is sufficiently short for the frequency response of the channel to be essentially constant over its operating bandwidth. In this case fading is said to be flat since all frequency components of a signal are subjected to the same fade at any given instant. When several or more propagation paths exist the fading of signal *amplitude* obeys Rayleigh statistics (due to the central limit theorem). If the spread of path delays is longer then the frequency response of the channel may change rapidly on a frequency scale comparable to signal bandwidth, Figure 12.28. In this case the fading is said to be frequency selective and the received signal is subject to severe amplitude and phase distortion. Adaptive equalisers may then be required to flatten and linearise the overall channel characteristics. The effects of flat fading can be combatted by increasing transmitter power whilst the effects of frequency selective fading cannot. For microwave links which are subject to flat fading a fade margin is usually designed into the link budget to offset the expected multipath (and rain induced) fades. The magnitude of this margin depends, of course, on the required availability of the link.

To reduce the necessary fade margin to acceptable levels diversity reception is sometimes employed. Figure 12.29(a) to (c) illustrates the principles of three types of diversity system, namely space (also called height), frequency, and angle diversity. In all cases the essential assumption is that it is unlikely that both main and diversity channel will suffer severe fades at the same instant. Selecting the channel with largest CNR, or combining channels with weightings in proportion to their CNRs, will clearly result in improved overall CNR.

12.5 Summary

Noise is present in all communications systems and, if interpreted to include interference, is always a limiting factor on their performance. Both thermal and shot noise have white

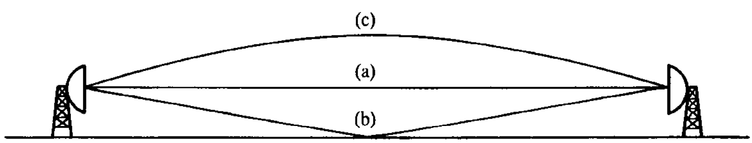

Figure 12.27 *Multipath in line of sight terrestial link due to: (a) direct path plus (b) ground reflection and/or (c) reflection from (or refraction through) a tropospheric layer.*

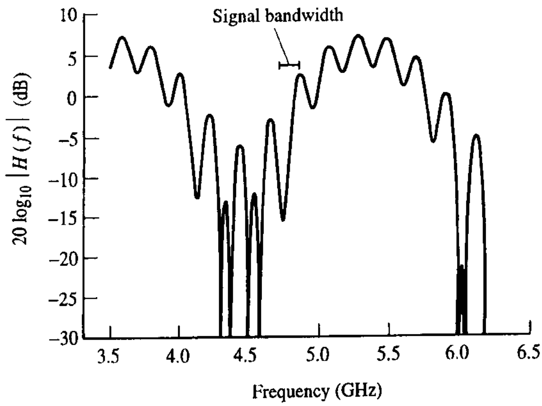

Figure 12.28 *Amplitude response of a frequency selective channel for 3-ray multipath propagation with ray amplitudes and delays of: 1.0, 0 ns; 0.9, 0.56 ns; 0.1, 4.7 ns.*

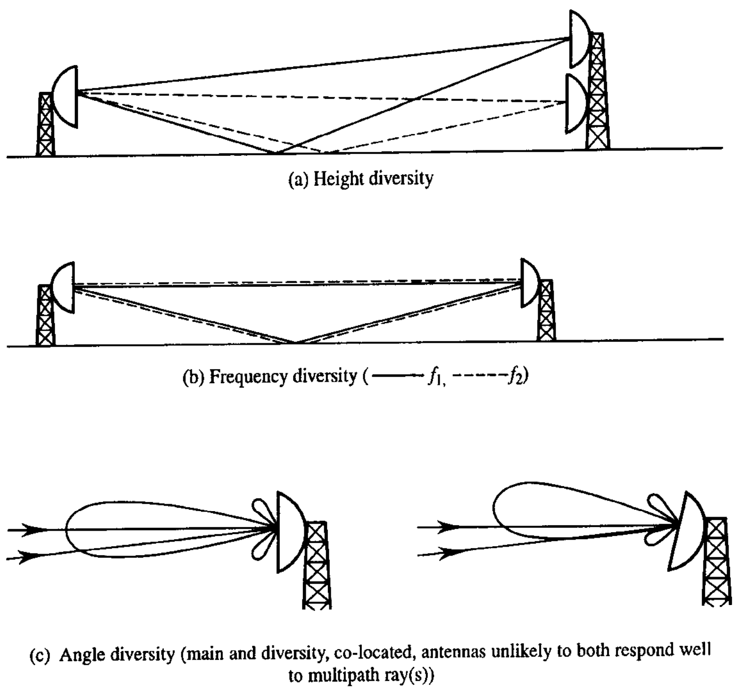

Figure 12.29 *Three types of diversity arrangements to combat multipath fading.*

power spectral densities and are therefore easy to quantify in a specified measurement bandwidth. In many systems calculations, non-white noise is often neglected and white noise is represented by an equivalent thermal spectral density or equivalent thermal noise temperature. The noise factor of a system is the ratio of its input SNR to output SNR when the system's source temperature is 290 K. The total noise available from several cascaded subsystems can be found using the noise factors, and gains, of each subsystem in the Friis formula.

Signal budgets are used to determine the carrier power present at the input to a communications receiver. In a radio system this involves antenna gains and path loss in addition to transmitter power. Two formulas for path loss are commonly used, namely the free space and plane earth formulas. Which is used for a particular application depends on frequency, antenna beamwidths, path clearance and the surface reflection coefficient of the earth beneath the path. Free space propagation results in a power density which decays as $1/R^2$ and a field strength which decays as $1/R$. Plane earth propagation results in a power density which decays as $1/R^4$ and a field strength which decays as $1/R^2$. For paths with particularly stable ground reflections it may be possible to mount antennas at an optimum height in order to take advantage of the 6 dB signal enhancement available at points of constructive interference.

Noise budgets in radio systems must account not only for the noise generated within the receiver itself but also for noise received by the antenna. The latter (assumed white) can be represented by an antenna temperature which is added to the equivalent noise temperature of the receiver to get the effective system input temperature. The effective system input noise power is then found using the Nyquist noise formula.

The link budget includes both signal and noise budgets and establishes the received carrier to noise ratio.

12.6 Problems

12.1. A preamplifier with 20 dB gain has an effective noise temperature of 200 K. What is its noise figure in dB? If this preamplifier is followed by a power amplifier of 30 dB gain with a noise figure F of 4 dB what is the degradation in dB for the overall noise figure of the two amplifier cascade compared to the preamplifier alone? [2.28 dB, 0.04 dB]

12.2. A microwave radio receiver has the following specification:

Antenna noise temperature	= 80K
Antenna feeder physical temperature	= 290 K, Loss = 2.0 dB
Low noise amplifier noise figure	= 2.2 dB, Gain = 12.0 dB
Frequency down converter conversion loss	= 6.0 dB
First IF amplifier noise temperature	= 870 K, Gain 30.0 dB
Second IF amplifier noise figure	= 25.0 dB, Gain 30.0 dB
Receiver bandwidth	=6.0 MHz

Assuming that the bandwidth of the receiver is determined by the second IF amplifier find the noise figure of the receiver, the system noise temperature of the receiver, and the noise power which you would expect to measure at the receiver output. [6.3 dB, 1025 K, –36.7 dBm]

12.3. A digital communications system requires a CNR of 18.0 dB at the receiver's detection circuit input in order to achieve a satisfactory BER. The actual CNR at the receiver input is 23.0 dB and the present (unsatisfactory) CNR at the detection circuit input is 16.0 dB. The source temperature of the receiver is 290 K. What noise figure specification must be met by an additional amplifier (to be placed at the front end of the receiver) for a satisfactory BER to be realised if the gain of the amplifier is: (a) 6.0 dB; (b) 10.0 dB; and (c) 14.0 dB? [3.3 dB, 4.4 dB, 4.8 dB]

12.4. A 36.2 GHz, 15.5 km, single hop microwave link has a transmitter power of 0 dBm. The transmit and receiver antennas are circularly symmetric paraboloidal reflectors with diameters of 0.7 m and 0.5 m, respectively. Assuming both antennas have ohmic efficiencies of 97% and aperture efficiencies of 67% calculate: (a) the transmitted boresight radiation intensity; (b) the spreading loss; (c) the free space path loss; (d) the effective isotropic radiated power; (e) the power density at the receiving antennas aperture; and (f) the received power. [3.64 W/steradian, 94.8 dB m^{-2}, 147.4 dB, 46.6 dBm, –78.2 dBW/m^{-2}, –57.1 dBm]

12.5. If the effective noise bandwidth of the receiver in Problem 12.4 is 100 kHz, the antenna aperture temperature is 200 K, its physical temperature is 290 K and the noise figure of the entire receiver (from antenna output to detector input) is 4.0 dB calculate the CNR at the detector input. [63.4 dB]

12.6. Derive the fundamental free space transmission loss equation using the concepts of spreading loss and antenna reciprocity. Identify the free space path loss in the equation you derive and explain why it is a function of frequency.

12.7. Explain the concept of path gain in the context of plane earth propagation. How does this lead to a $1/R^4$ dependence of power density? What is the equivalent distance dependence of field strength?

12.8. Show that equation (12.73) follows from equation (12.72) when $\rho e^{j\phi} = -1$.

12.9. A VHF communications link operates over a large lake at a carrier frequency of 52 MHz. The path length between transmitter and receiver is 18.6 km and the heights of the sites chosen for the location of the transmitter and receiver towers are 72.6 m and 95.2 m respectively. Assuming that the cost of building a tower increases at a rate greater than linearly with tower height find the most economic tower heights which will take full advantage of the 'ground' reflected ray. Estimate by what height the water level in the lake would have to rise before the 'ground' reflection advantage is lost completely. [80.3 m, 80.3 m, 69.0 m]

12.10. Digital MPSK transmissions carried on a 6 GHz terrestial link require a bandwidth of 24 MHz. The transmitter carrier power level is 10 W and the hop distance is 40 km. The antennas used each have 40 dB gain and filter, isolator and feeder losses of 4 dB. The receiver has a noise figure of 10 dB in the specified frequency band. Calculate the carrier-to-thermal-noise-power ratio at the receiver output. [62 dB]

Comment on any effects which might be expected to seriously degrade this carrier-to-noise ratio in a practical link. If a minimum carrier-to-noise ratio of 30 dB is required for the MPSK modulation what is the fade margin of the above link? [32 dB]

12.11. A QAM 4 GHz link requires a bandwidth of 15.8 MHz. The radiated carrier power is 10 W and the hop distance 80 km. It uses antennas each having 40 dB gain, has filter, isolator and antenna feeder losses totalling 7 dB at each end and a receiver noise figure of 10 dB. If the Rayleigh fade margin is 30 dB, calculate the carrier-to-thermal-noise-power ratio (CNR) at the receiver. What complexity of QAM modulation will this CNR support for an error rate of 10^{-6}? [32.5 dB, 128 state]

12.12. A small unmanned laboratory is established on the moon. It contains robotic equipment which receives instructions from earth via a 4 GHz communications link with a bandwidth of 30 kHz. The EIRP of the earth station is 10 kW and the diameter of the laboratory station's receiving antenna is 3.0 m. If the laboratory's receiving antenna were to transmit, 2% of its radiated power would illuminate the lunar surface and 50% would illuminate the earth. (The rest would be radiated into space which has the cosmic background brightness temperature of 3 K.) The noise figure of the laboratory's receiver is 2.0 dB. If the daytime physical temperature of the lunar surface is 375 K and the brightness temperature of the earth is 280 K estimate the daytime CNR at the laboratory receiver's detection circuit input. (Assume that the laboratory antenna has an ohmic efficiency of 98%, an aperture efficiency of 72%, and a physical temperature equal to that of the lunar surface. Also assume that the earth and moon both behave as black bodies, i.e. have emissivities of 1.0.) The distance from the earth to the moon is 3.844×10^8 m. [23.0 dB]

12.13. If the nighttime temperature of the lunar surface in Problem 12.12 is 125 K find the nighttime improvement in CNR over the daytime CNR. Is this improvement: (a) of real engineering importance; (b) measurable but not significant; or (c) undetectably small? [0.14 dB]