Chapter 9

Source and Receiver Noise

In the present chapter, we will discuss some of the fundamentals of the nature of optical sources and then go on to show that many of the same processes are at work in the electrical circuits which comprise receivers as well as the drivers of our optoelectronic sources. As it will turn out, the sources which are employed in transmitters really are noise sources. To impress an information stream on this noise signal, we must ingeniously try to gate on and off the noise emanating from this source. This stream of noise with time-varying mean value is then used to control the mean value of the circuit noise in a receiver circuit by using a detector as a gating element. We then try to reconstruct the root mean square noise value and hope that this somehow has something to do with the “deterministic” information stream we so ingeniously tried to impress on the source noise stream. One might wonder why we must take such a noisy picture of the world, or, further, if all the world is so noisy, why we would need to study specific types of noise.

That all matter around us is made up of fundamental building blocks called atoms that, at the electronic level at any rate, satisfy behavioral laws that we can write down on sheets of paper seems to indicate that the world around us is an eminently describable place. A fly in the ointment is, however, that the matter around us contains many, many atoms. Each atom can be in any of an uncountably infinite number of states depending on the available energy. That is, even at a fixed energy, a single atom may have more than one available state. Two atoms allowed to exchange energy (often called thermodynamic equilibrium) will have a multiplicity of states available at a fixed energy. Therefore agglomerations of atoms can be in multiplicities of states. Numbers of available states tend to grow as factorials of the numbers of atoms, and therefore these numbers of states tend to become astronomical even for barely macroscopic volumes. (The breakpoint between a macroscopic and microscopic volume is generally defined as a cube of one hundred angstroms on a side. Volumes from ten to a hundred angstroms on a side are generally referred to as mesoscopic.) Although the basic laws are simple (at least in the sense that we know what they are), practical knowledge of how a system will work requires manipulation of large amounts of data.

When we think about everyday information, we really think about pitifully small amounts, like the single bit contained in the statement, “go left, not right.” Even complicated experiments which require many channels of parallel data taking process amounts that are very small compared to the information required to specify the microscopic state of a chunk of chewing gum. The point being approached here really is that the difference between information and noise is not so much one of signal type as of amount. We can process a statement, “go left,” but if a warm chunk of material nudges us with fluctuating thermo-mechanical information about its state, we can’t process the quantity of information necessary to specify the quantum mechanical state of or, for that matter, even the entropy of the material, and therefore we call the sensory input noise. Noise is basically just huge quantities of information about the state of something we really don’t need to specify in any detail in order to go about our usual daily business.

The greatest content of information is carried by a signal which is totally random—that is, one in which the signal at one point in time and space is uncorrelated with that at another point. As the above discussion uncovered, noise is generally an extremely information-rich signal and therefore must appear to be nearly random. This is indeed the reason that noise analysis tends to lend itself to a probabilistic description. The description, however, will depend on the physical basis for the noise signal and therefore can take on a myriad of different mathematical forms. Some forms of noise can be related causally back to the quantum mechanical
uncertainty principle directly. Such processes will tend to exhibit Poisson or conditioned Poisson statistics. The uncertainty in the value of a state of the electromagnetic field is then due to the spontaneous rates of emission and absorption of photons. Other noise distributions are thermal in origin and tend to exhibit Gaussian statistics in terms of motion or negative binomial variate distributions in terms of their radiation characteristics. Still other forms of noise can owe their characteristics much more to their “propagation” than to their origin. Random processes can be strongly filtered by the dynamical processes that they trigger, such that their statistics vary depending on the time between their occurrence and observation. In other cases, reasonably simple events can trigger systems whose dynamical behavior is so complex that their output can contain more information than we can decipher about the process. In each of these cases, distinct statistical descriptions of the resultant noise can arise.

At the level of probability theory, there is a unique language with which to discuss noise processes. When it comes to practical descriptions of noise in terms of measurable system quantities, the metrics are naturally system-dependent. That is to say that the experimental problems of characterizing noise in a circuit versus a microwave cavity versus an optical cavity are different. In fact, these lead to whole different languages being used to describe noise in each of these fields.

In the discussion that follows, we will attempt to break up thermal noise into two subclasses (somewhat arbitrarily). One class will be spontaneous emission noise—that is, quantum noise due to random emission of photons from excited states. These states may be excited either by thermal excitation (section 9.1.1) or by an injected (9.1.2) or fed-back (9.1.3) signal. Section 9.1 includes a first subsection on blackbody radiation and then goes on in the next two subsections to show that the spectral lineshape of a poor coherence sub-threshold light source is the same as that of a laser biased well above threshold, although the linewidths may be orders of magnitude apart and they do not go smoothly into one another as the bias is changed. Laser jitter is broken off into a section of its own (section 9.2), being a non-equilibrium effect although still related to the properties of spontaneous emission. The second class will be that of circuit noise, which refers to the low-frequency (due to circuital filtering) fluctuation of voltages and currents that are due to coupling of circuits to the local ambient temperature in lossy components. We will point out why these types of noise really differ only in frequency content and not in cause—from the cavities exhibiting spontaneous emission noise. Section 9.3 contains detailed discussion of the generation of thermal circuit noise, shows that this noise is actually blackbody radiation in the lower-frequency portion of the thermal spectrum (section 9.3.1), then contains discussion of the concepts of noise temperature figure and noise. The $1/f$ noise of section 9.4 is another form of thermal noise but one that shows up in circuits and lasers as well as most everywhere else. Laser relative intensity noise gets a section of its own (9.5).

## 9.1 Some Basics of Optical Sources

Before turning the discussion to spontaneous emission and the myriad of effects it has on laser-like sources, it is perhaps a good idea to look at some of the inner workings of a few different types of sources at least qualitatively so that we have an anchor for the later, more involved considerations.

### 9.1.1 Some Thermal Sources

**A Star**

Figure 9.1 depicts the layers generally identified with a star which may be much like our own. The way a star generates radiation goes something like the following. There is a very dense core (Region 1), which is so dense that the pressure in it has ignited fusion reactions throughout. The core is essentially a large, continuously burning hydrogen bomb. As one goes out from the center of the star, however, the gravitational force will weaken, and therefore the density of material goes down. At some radius, the burning stops. Outside this radius (Region 2), the heat is still intense heat and the density is still very high. This region cannot obtain any form of thermodynamic equilibrium, and therefore the heat must be carried by convection. The region is referred to as the *convection zone*. At some further radius, local thermodynamic equilibrium is obtained. It is in this region (4) that the bubbling packets of heat from below will be converted to radiation. The density is still high enough in this region that the collision rate is so large as to cause the radiation to appear
to be a continuum even though it is coming from specific atomic transitions. There are many transitions contributing, and each one is greatly thermally broadened. In order to find the spectrum of this radiation in the next section of this chapter, we will need to consider what fields would exist in a cavity with blackbody walls. This region is known as the *photosphere*. Photosphere seems an apt name, as it is the region which generates light. The photosphere in essence looks like what we often refer to as a *blackbody radiator*. That is, it is a layer that we put in contact with a heat source from below. To remain in equilibrium, the photosphere must radiate all of the energy it takes in into the region above. This region above is defined by a density that is low enough that the interparticle collision rate becomes negligible. The temperature is low enough in this upper region that the thermal radiation is no longer significant. (The total thermal radiation is proportional to the fourth power of the temperature.) In this region (5) known as the *chromosphere*, radiation from the photosphere is absorbed in discrete lines. It is these absorption lines which tell astronomers on Earth what elements are present in the star. The corona (6) and solar wind are sufficiently rarified that they have little effect on the radiation stream.

### A Light Bulb

The main idea behind the operation of a light bulb is illustrated in Figure 9.2. The idea is much like that of a star or essentially any thermal source. A highly resistive filament in the pressurized bulb is heated by application of a voltage. The filament serves the purpose of the stellar core. The heating causes the pressurized gas about to speed up its average thermal velocity, thereby increasing the rate of collisions which in turns spreads out the heat, giving more and more gas molecules enough energy to make level transitions and thereby enabling these molecules to give up the energy by radiating optically. An arc lamp is a variant on this theme. In an arc, the resistive element is quite thick and has a gap. The heat in the gap is the most intense, and therefore most of the radiation comes from there. (Again, the total radiation is proportional to the fourth power of the temperature.) Therefore the radiation appears to emanate from something somewhat point-like. This makes it somewhat easier to couple such light into an optical system, as one could say that

![Figure 9.1: The generally accepted configuration for a regular star, where 1 is the nuclear burning zone, 2 is the convection zone, 3 is the photosphere, 4 is the chromosphere, 5 is the corona, and from the corona the stellar wind radiates out. The figure is not at all to scale.](image-url)
Figure 9.2: Illustration of the operation of a light bulb wherein an ac voltage applied to a resistive filament heats a gas sufficiently that it radiates.

the arc lamp has a much higher degree of spatial coherence than does the light bulb.

9.1.2 Lasers and LEDs

Laser Cavities

A typical schematic of a laser cavity is depicted in Figure 9.3. The gain medium can vary tremendously with laser type, but the use of the Fabry-Perot cavity is the common aspect. The mirrors need not be plane but can have a range of curvatures which satisfy a stability criterion. (See, for example, Yariv 1997 or Siegman 1986.) In fact, in ultra-high-power lasers (i.e. for laser fusion), sometimes both mirrors are taken to be concave, a clearly unstable situation but one which disperses the energy throughout the medium—that is, gives less probability of breaking down the medium. If the gain of the medium is high enough, the lost energy isn’t a problem.

Let’s consider how the modes of such a cavity might appear. If the mirrors were infinitely extended, perfectly flat, and perfectly reflecting and the medium was sufficiently rarified to be considered free space, then the modes would be transversely (i.e. in the \(x\) or \(y\) directions) polarized standing waves of the form

\[ E_N(z, t) = E_n \hat{e}_t \sin k_N z \cos \omega_n t, \]  

(9.1)

where

\[ k_n = \frac{n \pi}{L} \]

\[ \omega_n = k_n c \]  

(9.2)

and \( \hat{e}_t \) is a transverse (to \(z\)) unit vector. If the mirrors had curvature, then there would also be approximately transverse modes and, if one assumed the curvature were small enough (or the cavity long enough) that the curvature did not affect the longitudinal mode structure too much, then one could write
Figure 9.3: Schematic depiction of a laser cavity.

\[
E_{mn}(x, y, z, t) = E_n e^{i \psi_m(x, y)} \sin k_{nm} z \cos \omega_{nm} t.
\] (9.3)

If the mirrors focus on each other (a confocal resonator), then the modes are the Gauss-Hermite modes (Seigman 1986) that were found for the simple harmonic oscillator in Chapter 6. Generally, the spacing between transverse modes is much smaller than that between adjacent longitudinal modes. If a laser is run fully multimodal, then the longitudinal mode amplitudes will follow the shape of the gain curve and the Fourier spectrum, as is illustrated in Figure 9.4.

### Gas and Solid-State Lasers

As was discussed to some degree in Chapter 1, we will not pay too much attention to gas and/or solid-state lasers here, as in most cases in optical communications we would like to miniaturize things as much as possible. Even if one needs to use an external modulator, one would like to integrate it with a laser. Generally, gas and solid-state lasers have low gain per length, and therefore the mirrors really are part of an external cavity; they are not placed right in the medium. As the low gain requires that the mirrors are highly reflective (that the cavity be high-finesse), any high-speed modulation must be external, as it takes the time of multiple passes of light through the cavity in order to change the light output.

### Semiconductor Lasers

The semiconductor laser is based on the principle that, when an electron and hole recombine, energy is released. This energy can primarily be in the form of electromagnetic energy if the material and configuration of the material are chosen properly. Bandgaps and quasi Fermi levels are sketched for a direct and an indirect gap semiconductor in Figure 9.5. The point here is that photons carry energy but little momentum, whereas phonons carry little energy compared to their momentum. As we discussed previously, it has been assumed in the figure that the electrons rapidly come to an equilibrium in the conduction band and the holes in the valence band such that there are so-called Fermi levels near or in the conduction band and another near or in the valence band. Such levels never extend too deeply into a band, as the fact that they extend into a band at all indicates that there is a population inversion that will rapidly be eaten up by photon and phonon radiation. In the direct gap semiconductor, this inversion can be eaten up by pure photon radiation, as the minimum gap is at zero momentum. In the indirect semiconductor, both a photon and phonon must be emitted for a conduction band electron to relax to the valence band. Phonons are heat. An indirect gap semiconductor will burn up before it will lase. This is rather a shame, as both silicon and germanium have indirect gaps. It would be nice if one could make optical devices in the most common electronic material,
Figure 9.4: Possible spectrum of the field coming from a highly multimodal laser, where the central spikes represent the longitudinal modes corresponding to the lowest-order transverse mode and the smaller spikes around it represent high-order transverse modes. As laser cavities tend to be many wavelengths long, the central $n$ value, depending on laser type, may vary from hundreds to millions.

silicon. As it turns out, one must use binary alloys such as GaAs and GaN, ternary alloys such as GaAlAs or GaInAs, and quaternaries such as InGaAsP to make semiconductor lasers. A binary, not naturally occurring material such as GaAs will be much harder to process than pure, naturally occurring crystal such as silicon, but such is the situation.

The second requirement for lasing is a suitable cavity. The situation is as depicted in Figure 9.6. To present, the most common laser diodes are the so-called Fabry-Perot lasers. Vertical cavity surface-emitting lasers are also Fabry-Perot lasers but are called VCSELs and are differentiated from standard semiconductor lasers. A typical Fabry-Perot laser will have upper and lower contacts to inject current into the p-n junction and will use cleaved facets at either end of the junction to provide the feedback necessary for lasing. The VCSEL actually grows mirrors in the material system below the upper (transparent) contact and above the lower (transparent) contact. The configuration depicted in Figure 9.7 is indicative of the first semiconductor laser demonstrations in that the lasing region is depicted as a single p-n junction. The first lasers back in 1961 (Basov et al 1961, Hall et al 1962, Nathan et al 1962), however, could not be operated at room temperature due to high threshold currents. That is, as depicted in Figure 9.7, a typical power-versus-intensity curve for a laser diode appears much as a typical diode IV curve, but with one important difference. The V threshold ($V_{th}$) in an electrical diode is given by the bandgap voltage. In a laser diode, the threshold current $I_{th}$ is a complicated function of all the parameters of the device configuration and material parameters. Low threshold current requires enlightened design.

It was first Kroemer in 1964 who pointed out that a heterostructure would greatly improve the efficiency of recombination of electrons and holes in a semiconductor laser. The idea of the heterostructure was to grow slightly dissimilar materials on top of each other in order to trap carriers. The idea is depicted in Figure 9.8. The idea is that, if one can grow materials with different bandgaps on top of each other, then one can form a “trapping” region. By growing an intrinsic region (no doping thus little scattering) between regions of ternary Ga$_x$Al$_{1-x}$As, which has a higher bandgap, holes from the P region and electrons from the N region will flow into the I region and have a hard time getting out due to the potential barrier. The
Figure 9.5: Sketches of the band structure, quasi-levels for a pumped material, and path that an electron recombining must take, indicating that (a) in a direct semiconductor a single photon can be emitted in recombination, while in (b) a photon and a phonon must be emitted in order for recombination to occur.

Figure 9.6: Depiction of semiconductor laser configurations where (a) shows where the contacts and p-n layers are in a Fabry-Perot laser while (b) shows how the cavity is formed between the two end-cleaved crystal facets, and (c) shows the configuration of a VCSEL.
recombination probability goes up under these conditions. Further, a higher bandgap material will always exhibit a lower index of refraction (Pankove 1971). For this reason, the higher-index “active” region will act not only as a carrier trap but also as a waveguide for light. Indeed, it was just such a structure in which room-temperature semiconductor lasing was first demonstrated in 1970 (Alferov et al. 1970, Hayashi and Panish 1970). However, such a structure has no transverse (i.e. \( y \) direction) guidance and will therefore have a mode structure such as that sketched in Figure 8.4. From the point of view of dispersion, such a broad mode structure is not really desirable.

In order to guide in the transverse direction, one needs to perform more processing steps. Growing layers epitaxially one on top of another is, at this point in time, reasonably straightforward by either liquid-phase epitaxy (LPE), metalorganic chemical vapor deposition (MOCVD), or molecular beam epitaxy (MBE). Both MBE and MOCVD require high vacuum within the chamber but allow for much greater precision of growth than does LPE. Breaking the vacuum during processing can greatly lower the process yield. This was the reason that the purely planar structures such as those of Figure 9.9 were the first to be demonstrated. Any transverse guidance in these structures was so-called gain guidance. The idea there was to limit the width of the upper contact such that current was only injected into a narrow portion of the active region. The original index-guided structure was the ridge guide, followed by the buried heterostructure, both illustrated in Figure 9.9. The idea behind the ridge guide is to grow all the layers and then afterward to etch back a ridge, fill in the etch with SiO\(_2\), and then deposit the upper contact. The SiO\(_2\) serves both to block current and to lower the index of refraction outside the central guiding region in the GaAs. In such a structure, therefore, the light will tend to be guided in a two-dimensional waveguide located in the central region of the GaAs. The buried heterostructure, which requires reinsertion into the original growth chamber—that is, requires breaking and recreating vacuum conditions, does the same things as the ridge but much more strongly. That is, the \( N \) Ga\(_x\)Al\(_{1-x}\)As regions don’t just block, but completely reject, current from the upper electrode. As Ga\(_x\)Al\(_{1-x}\)As surrounds the active region in the buried heterostructure, the GaAs region is not just an effective guide but a real index guide in both dimensions. The transverse guidance should suppress the transverse mode structure of Figure 9.4, but one would expect the longitudinal modes to still exist. Interestingly enough, in the GaAs/Ga\(_x\)Al\(_{1-x}\)As system this is not the case. Transverse guidance leads to single-mode behavior. The GaAs/Ga\(_x\)Al\(_{1-x}\)As system, the first to be developed and which operates in the 0.85 \( \mu m \) wavelength region, is a special case, however. Single-mode fibers have minimum dispersion at 1.3 \( \mu m \) and minimum loss at 1.55 \( \mu m \), so as soon as the telecommunications industry looked to the use of single-mode fiber for long-haul transmission, there was a great push to make 1.3 \( \mu m \) and 1.55 \( \mu m \) sources. The InP/In\(_{1-x}\)Ga\(_x\)As\(_y\)P\(_{1-y}\) system, by variation of the \( x \) and \( y \), allows operation in either of these windows. In this material system, however, transverse guidance does not lead to single-mode operation, and further
Figure 9.8: Schematic depiction of (a) the layers grown to form a low-threshold semiconductor laser where the $P^+$ and $N^+$ layers are included for coupling to the contacts and (b) a plot of the minimum of the bandgap as a function of coordinate extending from the top contact through to the bottom contact of the semiconductor.
dispersive structures are necessary to force single-mode operation. The most common in use at present is the so-called distributed feedback structure, schematically depicted in Figure 9.10. The idea of these structures is that a long (in terms of wavelengths) grating can be quite spectrally selective, much more so than a Fabry-Perot cavity. By reflecting strongly at only one of the longitudinal mode wavelengths, the grating structure can pick out a single mode for lasing. A drawback of the DFB and DBR structures is that they do not have especially good direct modulation characteristics and therefore generally require external modulators.

The Interaction Between Radiation and an Atomic Medium

This sub-sub-section contains some of the material which is necessary for a more quantitative treatment of the semiconductor and/or atomic medium lasers. Some readers may want to forego this material on a first reading.

Figure 9.10: Sketches of (a) a distributed feedback (DFB) laser where the cladding region was etched with a sinusoidal pattern and then regrown with a slightly different material to cause a sinusoidal index variation in the channel and (b) a distributed Bragg reflector (DBR) structure where the sinusoidally modulated index region lies to either side of the active region.
CHAPTER 9. SOURCE AND RECEIVER NOISE

For a complete description of the interaction of a radiation field and an atomic or semiconductor medium we really need a coupled system of equations that describe the evolution of the field as well as the evolution of the polarization and inversion of the medium. In Chapter 7 (section 7.2), we presented (essentially without any derivation) a fully second-quantized (quantized field as well as matter) system of equations that would explain essentially anything about the light sources that we have discussed up to this point in this Chapter. We will rewrite these equations here for purposes of exposition for most readers but as well as for more quantitative use for the more advanced reader. The classical system which couples the time varying amplitude of the $i$th field mode, $b_i(t)$, with the polarization of the $k$th (for an atomic this $k$ is often replaced by an atomic coordinate $\mu$) electron momentum state, $p_k(t)$, and the relative inversion (magnitude of the difference in the conduction and valence band probabilities) for $k$th state, $d_k(t)$, has the fully quantum (discretized) realization (when considered as a stochastic process)

$$
\frac{db_i(t)}{dt} + \frac{(1 + i\omega_i \tau_{bi})}{\tau_{bi}} b_i(t) = -i \int_{-\infty}^{\infty} dk \rho_c(k) g^*_i(k)p_k(t) + F_i(t) \tag{9.4a}
$$

$$
\frac{dp_k(t)}{dt} + \frac{(1 + i\omega_{pk} \tau_{pk})}{\tau_{pk}} p_k(t) = i \sum_k g_{ik} b_i(t)d_k(t) + \hat{\Gamma}_{k-}(t) \tag{9.4b}
$$

$$
\frac{dd_k(t)}{dt} + \frac{d_k(t) - d^*_k}{\tau_{dk}} = 2i \sum_i (g^*_{ik} \hat{p}_i(t)\hat{b}_i(t) - g_{ik} \hat{p}^*_i(t)\hat{b}_i(t)) + \hat{\Gamma}_{kd}, \tag{9.4c}
$$

where $\hat{b}_i(t)$ ($\hat{b}^*_i(t)$) is the operator which absorbs (creates) a photon in the field mode labelled $i$; $\hat{p}_k(t)$ and $\hat{d}_k(t)$ are the operators which polarize and invert the $k$th atom in the medium, respectively, with $\hat{d}_eq$ the thermal “mean” operator of $\hat{d}_k$ where the $eq$ stands for equilibrium, meaning thermal equilibrium; $\omega_i$ and $\omega_{pk}$ are the angular frequencies of the field mode $i$ and the resonant transition angular frequency of the $k$th atom; $\tau_{bi}$, $\gamma_k$, and $1/\tau_{dk}$ are the damping constants for the $i$th field mode, where the $\tau_{bi}$ is also often expressed as the rate $\kappa_i = 1/\tau_{bi}$ and $k$th atomic dipole moment and $k$th atomic inversion; $g_{ik}$ is a coupling constant which gives the overlap between the $i$th field mode and $k$th atom; $\rho_c(k)$ is the density of states in the conduction band; and the $\hat{F}_i(t)$, $\hat{\Gamma}_{k-}(t)$, and $\hat{\Gamma}_{kd}(t)$ are zero-mean fluctuation operators which arise due to the absorption terms. The fluctuation operators indicate that, if something is absorbed, it will eventually return to the medium. The fluctuation operators satisfy the commutation relations

$$
\langle \hat{F}_i(t) \rangle = \langle \hat{F}^*_i(t) \rangle = \langle [\hat{F}^*_i(t), \hat{F}_i(t')] \rangle = \langle [\hat{F}_i(t), \hat{F}_i(t')] \rangle = 0 \tag{9.5a}
$$

$$
< [\hat{F}^*_i(t), \hat{F}_i(t')] > = 2\kappa_i \bar{m}(T) \delta(t - t') \tag{9.5b}
$$

$$
< [\hat{\Gamma}_{k-}(t), \hat{\Gamma}_{k+}(t')] > = 2\kappa_i \bar{m}(T + 1) \delta(t - t'), \tag{9.5c}
$$

$$
\langle \hat{\Gamma}_{k-}(t) \rangle = \langle \hat{\Gamma}_{k+}(t) \rangle = \langle \hat{\Gamma}_{kd}(t) \rangle = 0 \tag{9.6a}
$$

$$
[\hat{\Gamma}_{k+}(t), \hat{\Gamma}_{k'+}(t')^\dagger] = \left( \frac{1}{\tau_{pk}} (1 + \langle \hat{d}_k \rangle) + \frac{1}{2\tau_{eq}} (\hat{d}_eq - \langle \hat{d}_k \rangle) \right) \delta_{kk'} \delta(t - t') \tag{9.6b}
$$

$$
[\hat{\Gamma}_{k-}(t), \hat{\Gamma}_{k'+}(t')^\dagger] = \left( \frac{1}{\tau_{pk}} (1 - \langle \hat{d}_k \rangle) - \frac{1}{2\tau_{eq}} (\hat{d}_eq - \langle \hat{d}_k \rangle) \right) \delta_{kk'} \delta(t - t') \tag{9.6c}
$$

$$
[\hat{\Gamma}_{kd}(t), \hat{\Gamma}_{k'd'}(t')] = \frac{2}{\tau_{dk}} (1 - \hat{d}_eq \langle \hat{d}_k \rangle) \delta_{kk'} \delta(t - t') \tag{9.6d}
$$

$$
[\hat{\Gamma}_{k+}(t), \hat{\Gamma}_{kd}(t')] = \frac{1}{\tau_{dk}} (1 + \hat{d}_eq \langle \hat{p}_k \rangle) \delta_{kk'} \delta(t - t') \tag{9.6e}
$$

$$
[\hat{\Gamma}_{k-}(t), \hat{\Gamma}_{kd}(t')] = \frac{1}{\tau_{dk}} (1 - \hat{d}_eq \langle \hat{p}_k \rangle) \delta_{kk'} \delta(t - t'), \tag{9.6f}
$$

and

$$
\langle \hat{\Gamma}_j(t) \hat{F}_i(t') \rangle = \langle \hat{\Gamma}_j(t) \rangle \langle \hat{F}_i(t') \rangle = 0 \tag{9.7a}
$$
CHAPTER 9. SOURCE AND RECEIVER NOISE

9.4 In order to self-consistently describe laser noise, these equations take the form

\[ \langle \hat{\Gamma}_j(t)\hat{F}_i(t') \rangle = 0. \]  

(9.7b)

where the \( \bar{n}(T) \) is the average thermal mode occupancy, that is, the average number of thermal photons within a mode, and it is also given by a relation to the equilibrium temperature

\[ \bar{n}(T) = \frac{1}{e^{\frac{k_B T}{kT}} - 1} \]

where the \( T \) is the temperature and \( k_B \) is Boltzmann’s constant.

As was discussed in the last chapter, we don’t generally need a fully quantized set of equations unless we are really discussing the generation of light. By using conditional Poisson counting at the detector, we can include what is generally in optical communications the most important quantum mechanical noise source—that which is called shot noise. To reduce the above to the system to its semiclassical form, we need to take expectation values, i.e.

\[ \langle \hat{b}_i \rangle = b_i, \]

(9.8)

where the angular bracket is an expectation and therefore the \( b_i \) is now a simple function of time. There is a problem with taking expectations of products of operators since, being in a coupled system of equations, they may be correlated. However, such approximations as

\[ \langle \hat{p}_k \hat{b}_i^\dagger \rangle \approx p_k b_i^* \]

(9.9)

are the basis of the semiclassical picture. Taking expectations and using relations such as (9.9), we find that we can write

\[ \frac{db_i(t)}{dt} + \frac{(1 + i\omega_i \tau_{bi})}{\tau_{bi}} b_i(t) = -i \sum_k g_{ik}^* p_k(t) \]  

(9.10a)

\[ \frac{dp_k(t)}{dt} + \frac{(1 + i\omega_{jk} \tau_{pk})}{\tau_{pk}} p_k(t) = i \sum_i g_{ik} b_i(t) \]  

(9.10b)

\[ \frac{dd_k(t)}{dt} + \frac{d_k(t) - d_{eq}}{\tau_{dk}} = 2i \sum_j (g_{jk}^* p_k(t) b_j^*(t) - g_{jk} p_k(t) b_j(t)), \]

(9.10c)

where the \( b_i(t), p_k(t), \) and \( d_k(t) \) are now simply functions of time. The \( d_{eq} \) now represents the pump. If the system were in thermal equilibrium,

\[ d_{eq} = \frac{1 - e^{\beta \omega_i}}{1 + e^{\beta \omega_i/kT}}, \]

(9.11)

sometimes a laser is described as a system in equilibrium at a negative temperature. The system of (9.10) was derived semiclassically in section 7.1. It should be noted, however, that in 7.2 it was pointed out that, although the fluctuating operators don’t formally show up in (9.10), this is a failure of the semiclassical picture. There really should be fluctuating functions in (9.10) which satisfy averages analogous to those in 9.4? in order to self-consistently describe laser noise. These equations take the form

\[ \frac{db_i(t)}{dt} + \frac{(1 + i\omega_i \tau_{bi})}{\tau_{bi}} b_i(t) = -i \int_0^\infty dk \int_0^\infty dp e(k g_{ik}^* p_k(t) + f_{bi}(t)) \]

(9.12a)

\[ \frac{dp_k(t)}{dt} + \frac{(1 + i\omega_{jk} \tau_{pk})}{\tau_{pk}} p_k(t) = i \sum_i g_{ik} b_i(t) + f_{p}(t) \]  

(9.12b)

\[ \frac{dd_k(t)}{dt} + \frac{d_k(t) - d_{eq}}{\tau_{dk}} = 2i \sum_j (g_{jk}^* b_j^*(t) p_k(t) - g_{jk} b_j(t) p_k(t)) + f_d(t), \]

(9.12c)

where each of the Langevin source terms \( f_i(t) \) satisfy

\[ \langle f_{bi}(t)f_{bi}^*(t') \rangle = 2\kappa_i \bar{n} \delta_{ij} \delta(t - t') \]

(9.13a)

\[ \langle f_{p}(t)f_{p}^*(t') \rangle = -2 \frac{d_k}{\tau_{pk}} + \frac{d_k - d_k}{\tau_{dk}} \delta_{kk} \delta(t - t') \]

(9.13b)
\[
\langle f_d(t)f_d(t') \rangle = 2 \frac{1 - d_{eq}dk}{\tau_d} \delta_{kk'} \delta(t - t') \tag{9.13c}
\]

\[
\langle f_p(t)f_d(t') \rangle = -2 \frac{d_{eq}}{\tau_d} p_k \delta_{kk'} \delta(t - t') \tag{9.13d}
\]

where the rate \( \kappa_i \) has replaced the inverse photon lifetime \( \tau_{bi} \) in the above source term relations.

9.2 Spontaneous Emission Noise

9.2.1 Modes in a Cavity

As is well-known, heating a material system to a high enough temperature is a means to generate light, as is evidenced by campfires as well as incandescent lights. Much of this section will have to do with how such sources work. Interestingly enough, their understanding requires a more complete quantum mechanical picture than does, for example, understanding laser action. (It is fortunate that neither Edison nor the early cave men were aware of this, or we might still be in the dark.)

Consider the cavity depicted in Figure 9.11. In such a cavity, we could express Maxwell’s equations in the free space form (see sections 2.2, 3.1, 4.1, 5.2, and 6.3 as well as Papas 1965 or Mickelson 1992)

\[
\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{9.14a}
\]

\[
\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} \tag{9.14b}
\]

\[
\mathbf{D} = \epsilon_0 \mathbf{E} \tag{9.14c}
\]

\[
\mathbf{B} = \mu_0 \mathbf{H} \tag{9.14d}
\]

subject to the boundary conditions that

\[
dA \times \mathbf{E} = 0 \text{ on } A \tag{9.15a}
\]

\[
dA \times \mathbf{H} = 0 \text{ on } A; \tag{9.15b}
\]

that is, the tangential \( \mathbf{E} \) and \( \mathbf{H} \) fields disappear on the walls. We are aware that certain shapes at certain frequencies will satisfy the equations and boundary conditions. The eigenmodes \( \mathbf{E}_n(x, y, z, t) \) and \( \mathbf{H}_n(x, y, z, t) \), that is, the combination of shape and time dependence which can be associated with frequencies that we call eigenfrequencies \( \omega_n \), and can be expressed in the time and shape combinations

\[
\mathbf{E}_n(x, y, z, t) = \Re \left\{ u_n(x, y, z) e^{-i\omega_n t} \right\} \tag{9.16a}
\]

\[
\mathbf{H}_n(x, y, z, t) = -\frac{1}{i\omega_n \mu_0} \nabla \times \mathbf{E}_n(x, y, z, t). \tag{9.16b}
\]

That is, when the shapes are real functions then they are constant in time and the whole eigenfield just sinusoidally oscillates with time. If the shapes \( u_n(\mathbf{r}) \) (where the vector \( \mathbf{r} = (x, y, z) \)) are complex, then they may also perform such motion as rotation with time. The total field in the cavity is then a sum of all of these modes, that is shapes together with their time dependencies, but with a set of time dependent coefficients, which allow the cavity fields to satisfy a set of time boundary conditions, be they initial conditions, or driven (ongoing excitation) conditions. The total field in such a cavity is then expressible in terms of a set of excitation coefficients \( B_n(t) \) such that

\[
\mathbf{E}(x, y, z, t) = \Re \left[ \sum_n B_n u_n(x, y, z) e^{-i\omega_n t} \right]. \tag{9.17}
\]
CHAPTER 9. SOURCE AND RECEIVER NOISE

Figure 9.11: Depiction of a cavity with highly conductive walls and a homogeneous interior with permittivity $\varepsilon_0$ and permeability $\mu_0$.

We are really interested in the distribution of energy in the cavity. The energy in the cavity can be found from the use of Poynting’s theorem (see Chapter 4),

$$\int \mathbf{S} \cdot d\mathbf{A} = -\frac{\partial}{\partial t} \int_{V} \left( \frac{\mu_0}{2} \mathbf{H} \cdot \mathbf{H} + \frac{\varepsilon_0}{2} \mathbf{E} \cdot \mathbf{E} \right) dV,$$

(9.18)

where $V$ is the volume enclosed by the closed surface $A$ and where $\mathbf{S}$ is defined by

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}.$$

(9.19)

The left-hand side of (9.18) represents the power crossing the closed surface $A$.

To simplify the argument somewhat, let’s consider that the cavity is many wavelengths across and therefore, except near the cavity walls, the modes are plane wave-like. In such a case, the fields $\mathbf{E}_n$ and $\mathbf{H}_n$ will essentially be transverse, and further,

$$\mathbf{E}_n = \eta_0 \hat{e}_s \times \mathbf{H}_n,$$

(9.20)

where $\eta_0$ is the free-space impedance

$$\eta_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}}$$

(9.21)

and $\hat{e}_s$ is a unit vector in the direction of $\mathbf{S}$ such that

$$\hat{e}_s = \frac{\mathbf{S}}{|\mathbf{S}|}.$$  

(9.4)

With this, we can write that the electromagnetic power leaving a closed surface $A$, $P_{em}$, which is also equal to $\int \mathbf{S} \cdot d\mathbf{A}$, can be expressed as

$$P_{em} = -\frac{\partial}{\partial t} \int_{V} \varepsilon_0 \mathbf{E} \cdot \mathbf{E} dV.$$

(9.23)

However, the power leaving the volume must be equal to the negative of the time derivative of the potential energy $V(\mathbf{E})$ in the volume such that

$$P_{em} = -\frac{\partial}{\partial t} V(\mathbf{E}),$$

(9.24)

leaving us with

$$V(\mathbf{E}) = \varepsilon_0 \int \mathbf{E} \cdot \mathbf{E} dV.$$  

(9.25)
If we normalize our vector shapes \( \mathbf{u}_n(x, y, z) \) such that
\[
\int \mathbf{u}_n(x, y, z) \cdot \mathbf{u}_m(x, y, z) \, dr = \delta_{nm},
\] (9.26)
then we find that
\[
V(\mathbf{E}_n) = V(B_n) = \epsilon_0 B_n^2.
\] (9.27)

There is another possible mode normalization that is the one that we shall actually find to be more convenient to use. It is the photon normalization and it is the one that is generally used when interacting radiation with matter, or, for example, expressing integer conditioning numbers for use in discrete probability density functions. In this picture, we express the electric field in the form
\[
\mathbf{E}(\mathbf{r}, t) = -i \sum_i [b_i(t) - b_i^*(t)] \mathbf{u}_n(\mathbf{r}).
\]

If the normalization integral is then defined by
\[
\int d^3 \mathbf{r} \, \mathbf{u}_i(\mathbf{r}) \mathbf{u}_{i'}(\mathbf{r}) = \frac{\hbar \omega_i}{2\epsilon_0} \delta_{ii'}
\]
then the potential energy function \( V(\mathbf{E}) \) for the \( i \)th takes the form
\[
V(\mathbf{E}) = |b_i(t)|^2 \hbar \omega_i
\]
which is to say that we can make the identification that
\[
|b_i(t)|^2 = m_i(t)
\]
is the number of photons in the \( i \)th mode during the temporal mode occurring at the time \( t \). The amplitude \( b_i(t) \) in this normalization is called the photon amplitude, and the normalization, as we mentioned above, is often called the photon normalization.

In analogy with the case of Newtonian mechanics (or quantum mechanics, for that matter), we can define a conserved total energy \( E_{\text{tot}} \) consisting of kinetic plus potential energy \( E_{\text{kin}} + V(x) \). As is discussed in Chapter 6, a way to derive Newtonian mechanics from energy conservation is to write that
\[
E_{\text{tot}} = E_{\text{kin}} + V(x),
\] (9.28)
where \( E_{\text{kin}} \) is given by
\[
E_{\text{kin}} = \frac{1}{2} m \left( \frac{\partial x}{\partial t} \right)^2,
\] (9.29)
where \( m \) is a mass or, in the general case, a unit-preserving constant. We then know that the total energy is conserved or that
\[
\frac{\partial}{\partial t} E_{\text{tot}} = 0,
\] (9.30)
which can be rewritten
\[
\frac{\partial}{\partial t} \left[ \frac{1}{2} m \left( \frac{\partial x}{\partial t} \right)^2 + V(t) \right] = 0,
\] (9.31)
which has solutions
\[
m \frac{\partial^2 x}{\partial t^2} + \frac{\partial V}{\partial x} = 0
\] (9.32a)
\[
\frac{\partial x}{\partial t} = 0.
\] (9.32b)

In our electromagnetic field case, then, we can make the identification that
\[
E_{\text{kin}} = \frac{\epsilon_0}{\omega_n^2} \left( \frac{\partial B_n}{\partial t} \right)^2,
\] (9.33)
where \( m \) has been replaced by the constant \( \epsilon_0/\omega_n^2 \) to preserve dimensions, and then write

\[
\frac{\partial}{\partial t} \left[ \frac{\epsilon_0}{\omega_n^2} \left( \frac{\partial B_n}{\partial t} \right)^2 + GB_n^2 \right] = 0,
\]

which has solutions

\[
\frac{\partial^2 B_n}{\partial t^2} + \omega_n^2 B_n = 0 \tag{9.35a}
\]

\[
\frac{\partial B_n}{\partial t} = 0. \tag{9.35b}
\]

The second solution (b) is not very interesting. We know that the \( B_n(t) \) coefficients need to have a time dependence as we have already absorbed a factor of \( e^{-i\omega t} \) into them. If they have no time derivative, then they must be just zero. The first solution is the one of interest and happens to be the equation of motion of a harmonic oscillator. In a material, any atom or molecule for small-amplitude motion will exhibit the motion of a harmonic oscillator. (See Chapter 6.) This is the reason that thermodynamicists have spent so much effort on finding the properties of harmonic oscillators interacting with thermal reservoirs.

It can be of interest to compare the above with the photon normalized version of these equations. From separation of variables in the curl of \( \mathbf{E}(\mathbf{r}) \) equation, we can find that

\[
\nabla \times \nabla \times \mathbf{u}_i(\mathbf{r}) + \frac{\omega_i^2}{c^2} \mathbf{u}_i(\mathbf{r}) = 0.
\]

When the \( \mathbf{E}(\mathbf{r}, t) \) is then expanded in terms of \( b_i(t) \), the time dependent piece of the separation then satisfies

\[
\frac{db_i(t)}{dt} + i\omega_i b_i(t) = 0,
\]

and

\[
\frac{db_i^*(t)}{dt} - i\omega_i b_i^*(t) = 0.
\]

If one then defines quadrature coefficients (where the \( i \) subscript will be suppressed due to the \( c \) and \( s \) subscripts, hopefully for the present without any confusion) by

\[
b_c(t) = b_c(t) + ib_s(t),
\]

\[
b_c^*(t) = b_c(t) - ib_s(t),
\]

\[
b_c(t) = \frac{b_c(t) + b_c^*(t)}{2},
\]

\[
b_c(t) = \frac{b_c(t) - b_c^*(t)}{2i}.
\]

One then notes that the \( b_c(t) \), and \( b_s(t) \) now satisfy the harmonic oscillator equation.

### 9.2.2 Blackbody Radiation

Let’s now say that the walls of the cavity of Figure 9.11 are no longer perfectly conducting but perfectly absorbing as well as perfectly re-emitting and in contact with a large reservoir which is in thermal equilibrium at a temperature \( T \). (No material can really have these ideal “blackbody” properties, but some come close.) Further, such a “black” enclosure will still force the tangential fields to be approximately zero on the surface, so the modes should not be significantly changed from their perfectly conducting wall condition. In such a case, the modes (harmonic oscillators) of the field in the cavity will be continuously exchanging energy with the external energy reservoir. The potential energy will no longer be a constant. Indeed, this is the case we treated in detail back in chapter 7 when discussion the Langevin equations. One effect of the wall will be to add a damping term to each mode. Energy from a mode that is absorbed into the wall is then redistributed back into cavity modes through radiation emitted by the wall back into cavity modes. As the
wall is heated thermally as well as radiatively, and comes to equilibrium thermally as well as radiatively, the radiation re-emitted back into the cavity from the wall is generally not the same one from which the energy was originally extracted, and in fact, the number of photon emissions by the wall is not going to be at all the same as the number of absorptions, except perhaps in some form of a long time average. Of course, when steady state is achieved, the average amount of energy redistributed into a mode will be equal to the average amount of energy absorbed by the wall from that mode. For example, if originally, only one cavity mode is excited, eventually all of the cavity modes will be excited and each one in just the right amount such that the distribution of radiation in the cavity is the black body (Planck) distribution. That is to say that some kind of ensemble average over the potential energy of each mode will be constant in the mean in steady state despite what may be significant fluctuation from this mean. This constancy of average mode energy puts a constraint on the Langevin fluctuation noise source $f_{bi}(t)$ in the relation for the photon normalized mode amplitude $b_i(t)$

$$\frac{db_i(t)}{dt} + \frac{(1 + i\omega_i T_{bi})}{\tau_{bi}} b_i(t) = f_{bi}(t)$$

such that we can also write that

$$\langle f_{bi}(t) \rangle = 0$$

and

$$\langle f_{bi}(t) f_{bi}^*(t') \rangle = 2\kappa_i \frac{1}{\exp \left( \frac{h\omega}{kT} \right) - 1}$$

where $\kappa_i = 1/\tau_{bi}$ is the rate of photon absorption that is naturally given by the inverse of the photon lifetime in the cavity mode $i$, $\tau_{bi}$ and the average number of photons in a spatial-temporal mode $\bar{m}$ is given by the last term of the second order correlation function, namely,

$$\bar{m} = \frac{1}{\exp \left( \frac{h\omega}{kT} \right) - 1}.$$ 

Boltzmann has told us that, at a temperature $T$, the probability that a state of a harmonic oscillator in contact with a thermal reservoir at temperature $T$ with potential energy $\mathcal{E}$ is occupied will be given by

$$P(\mathcal{E}) = \frac{e^{-\mathcal{E}/kT}}{\sum_{i=\text{all possible states}} P(\mathcal{E}_i)} = \frac{e^{-\mathcal{E}/kT}}{Z}, \quad (9.36)$$

where the second equality sign can be used to define Boltzmann’s partition function $Z$. To conform to standard interpretation, we now note that a photon has an energy $\bar{h}\omega$ ($\hbar\omega = h\nu$) to allow us to talk about the $\epsilon_0 B_n^2$ being a discretized quantity. Then for one of our simple harmonic oscillators, we can evaluate $Z$ as follows:

$$Z = e^{-\hbar\omega/2kT} \sum_{n} e^{-n\hbar\omega/kT}, \quad (9.37)$$

where now we can use the identity

$$\sum_{x} x^n = \frac{1}{1 - x} \quad (9.38)$$

to obtain

$$Z = e^{-\hbar\omega/2kT} \frac{1}{1 - e^{-\hbar\omega/kT}}. \quad (9.39)$$

Multiplying top and bottom by $e^{\hbar\omega/kT}$, we obtain the final form

$$Z = e^{-\hbar\omega/2kT} \frac{e^{\hbar\omega/kT}}{e^{\hbar\omega/kT} - 1}. \quad (9.40)$$
Using the partition function of (9.40) gives that the probability of the cavity being in a state with \(s\) photons is

\[
P(m\hbar\omega) = e^{-m\hbar\omega/kT} \frac{e^{\hbar\omega/kT}}{e^{\hbar\omega/kT} - 1}.
\]  

(9.41)

The mean number of photons in a mode at a temperature \(T\) must therefore be

\[
\bar{m} = \langle m \rangle = \frac{e^{-\hbar\omega/2kT}}{Z} \sum_m m e^{-m\hbar\omega/kT},
\]  

(9.42)

where we note that the term in the sum is expressible as a derivative of \(x = \hbar\omega/kT\), which gives us the form

\[
\bar{m} = \langle m \rangle = -\frac{e^{-\hbar\omega/2kT}}{Z} \frac{\partial}{\partial x} \sum_m e^{-mx}.
\]  

(9.43)

If we now again use the identity of (9.38), we obtain

\[
\bar{m} = \langle m \rangle = -\frac{e^{-\hbar\omega/2kT}}{Z} \frac{\partial}{\partial x} \left( \frac{1}{1 - e^{-x}} \right).
\]  

(9.44)

The derivative can be performed analytically to obtain

\[
\bar{m} = \langle m \rangle = \frac{e^{-\hbar\omega/2kT}}{Z} \frac{e^{-\hbar\omega/kT}}{(1 - e^{-\hbar\omega/kT})^2}.
\]  

(9.45)

Using the definition of the partition function then gives us the desired form of the occupancy law for bosons:

\[
\bar{m} = \langle m \rangle = \frac{1}{e^{\hbar\omega/kT} - 1},
\]  

(9.46)

where the last step used (9.40).

We are now ready to complete a derivation of the Planck distribution. To find an expression for the energy per unit volume per unit frequency at frequency \(\nu = \omega/2\pi\) within the cavity \(\rho(\nu)\), we need to evaluate the following:

\[
\rho(\nu) d\nu = \left( \frac{\text{modes}}{\text{volume}} \right) \left( \frac{\text{avg energy}}{\text{mode freq}} \right).
\]  

(9.47)

Clearly, the electromagnetic field has two independent polarizations. From the above calculation, we know that

\[
\text{avg energy} = \bar{m}\hbar\omega,
\]  

(9.48)

and we also know that the density of states in a uniform \(k\)-space is \(d^3k\), which we need to normalize by \((2\pi)^3\) to obtain:

\[
\rho(\nu) d\nu = 2 \text{polarizations} \times \frac{d^3k}{(2\pi)^3} \bar{m} \hbar\nu,
\]  

(9.49)

which we can algebraically simplify to

\[
\rho(\nu) d\nu = 4 \frac{\hbar\nu^3 d\nu}{e^{\hbar\nu/kT} - 1},
\]  

(9.50)

which, when divided on both sides by \(d\nu\), is just Planck’s radiation formula. Note that the \(d^3k/(2\pi)^3\) assumes a large, uniform cavity, an approximation we have already used to obtain the harmonic oscillator form of the modal coefficient equation. Quantum confinement can lead to quite different densities of states and can even be used to modify such quantities as the spontaneous emission rate. Also note that the assumption that the average energy is \(\langle s \rangle \hbar\nu\) has ignored a zero-point energy \(\hbar\nu/2\) (see Chapter 6), which we will need to put back in at a later time.

We now wish to use Planck’s formula along with some simple considerations in order to calculate the rate at which spontaneous emission is generated in a cavity in which both matter and radiation are confined.
Figure 9.12: The three processes which occur in the semiclassical theory of the interaction of radiation and matter to bring the composite system to equilibrium: (a) stimulated absorption, (b) spontaneous emission, and (c) stimulated emission.

Let’s assume now that, along with the radiation in the cavity, we put a number of two-level atoms whose energy levels are separated by an energy given by

\[ E = \hbar \omega = h \nu. \]  \hfill (9.51)

In equilibrium, we are aware that the occupancies \( N_1 \) and \( N_2 \) of the two atomic states should satisfy Boltzmann’s relation

\[ \frac{N_2}{N_1} = e^{-h\nu/kT}. \]  \hfill (9.52)

To get more considerations, we need to consider in a bit more detail what processes can take place to bring (and keep) the matter in equilibrium with the radiation which is already assumed to be in equilibrium with the walls, which are at a temperature \( T \).

Consider the diagrams in Figure 9.12. These are the processes which couple the radiation field to the atomic states and thereby allow these two subsystems to come to equilibrium. Clearly, (a) and (c) are stimulated processes in that the field itself stimulates their occurrence, and therefore the process has memory of the state of the field in the cavity. Process (b), spontaneous emission, simply happens independent of what is occurring in the cavity and in that sense is a truly pure noise source.

Consider Figure 9.13. This is a schematic depiction of the end result of the three processes of Figure 9.12. For the radiation and matter to be in equilibrium, clearly we will need that

\[ W_{12}N_1 = W_{21}N_2. \]  \hfill (9.53)

That is, the upcoupling rate times the amount of stuff to upcouple needs to be equal to the downcoupling rate times the amount of stuff to downcouple at equilibrium. We can express

\[ W_{12} = B_{12}\rho(\nu) \]  \hfill (9.54a)

\[ W_{21} = B_{21}\rho(\nu) + A_{21}, \]  \hfill (9.54b)
where the $A$ and $B$s are Einstein’s $A$ and $B$ coefficients, with $B_{12}$ being the stimulated absorption coefficient, $B_{21}$ the stimulated emission coefficient, and $A_{21}$ the spontaneous emission rate. Note that the stimulated coefficients appear only with radiation density terms. Using this in (9.48) yields

$$B_{12} \rho(\nu) N_1 = (B_{21} \rho(\nu) + A_{21}) N_2,$$

which can be rewritten in the form

$$\rho(\nu) \left[ \frac{B_{12} N_1 - B_{21} N_2}{N_2} \right] = A_{21}. \quad (9.56)$$

Using (9.45) and (9.47) in (9.56) yields

$$\frac{4h\nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1} (B_{12} e^{h\nu/kT} - B_{21}) = A_{21}. \quad (9.57)$$

Basically, (9.57) is one equation in three unknowns, which seems to pose a problem. However, $A_{21}$, $B_{12}$, and $B_{21}$ are all atomic coefficients and therefore must be temperature-independent. In order for this to be true, we see that we need

$$B_{21} = B_{12}, \quad (9.58)$$

which yields

$$A_{21} = \frac{4h\nu^3}{c^3} B_{21}. \quad (9.59)$$

We can actually do even better in finding the spontaneous rate than (9.59). Recall that earlier (section 8.1) we discussed a coupled radiation and atomic system during the discussion of Fermi’s golden rule. In a two-level system coupled to radiation, we recall that the wave function of an atomic valence electron interacting with radiation which is not too far detuned from a specific $1 \rightarrow 2$ level transition could be expressed as

$$\psi(r, t) = a_1(t) \psi_1(x, y, z) e^{-i\omega_1 t} + a_2(t) \psi_2(x, y, z) e^{-i\omega_2 t}. \quad (9.60)$$

Further, in our present case, the time rate of change of the probability of remaining in the lower state must be given by

$$\frac{\partial}{\partial t} (a_1^* a_1) = B_{12} \rho(\nu). \quad (9.61)$$

Using Schrödinger’s equation for a two-level system together with the $\psi(r, t)$ of (9.60) can yield that

$$B_{12} = \frac{2\pi^3}{3} \frac{\mathcal{P}^2}{h^2 \epsilon_0}, \quad (9.62)$$

where $\epsilon_0$ is the dielectric constant of free space and $\mathcal{P}^2$ is the square of the dipole matrix element

$$\mathcal{P} = -e \int \psi_1(r) \psi_2^*(r) d^3r, \quad (9.63)$$
Figure 9.14: Schematic depictions of (a) a classical dipole; (b) a quantum mechanical dipole; and (c) and (d) the two lowest-order wave functions of some atom.

and therefore the spontaneous emission rate can be given by

$$A_{21} = \frac{8\pi^3}{3} \frac{\nu^3}{\hbar \epsilon_0} P^2.$$  \hspace{1cm} (9.64)

As was mentioned earlier, the spontaneous as well as the stimulated rates are atomic quantities and therefore must depend on the atoms taking part in an interaction. That is why the $P$ had to be there. One might recall that a usual classical dipole moment is defined by

$$p = -e r,$$  \hspace{1cm} (9.65)

where $r$ is the radius vector defined in Figure 9.14(a). A separation of a plus and minus charge gives rise to a dipolar field. In Figure 9.14(b), an interpretation of a quantum dipole is given. Here the electron is drawn as a cloud, and an integral over this cloud will determine an expectation value of a radius vector which can be used in the expression of (9.65). In reality, that wave function in Figure 9.14(b) is really a superposition of wave functions which may appear as those depicted in Figures 9.14(c) and (d). The antisymmetry of the wave function depicted in (d) will cause the electron wave function to have a time-varying, averaged radial position that will oscillate back and forth about the positive nucleus at a beat frequency given by the energy difference between the states divided by $\hbar$. This indeed is what equation (9.63) is telling us.

Interestingly enough, equation (9.63) indicates that the way to optimize the value of $P$ is to have states that have opposite parity—that is, $r$ is an odd function, so the integral will be maximized when the product of $\psi_1$ and $\psi_2$ is also an odd function, which is optimally achieved by having one of the functions odd and one even. This is related to the so-called dipole selection rules, which state that dipole transitions only occur between states of opposite parity. (It should be noted here that, for sufficiently large and therefore complex atoms, these selection rules don’t work, but this is actually due to the assumption that valence states of large atoms still have ideal parity, which they do not, basically due to spin-spin terms in the Hamiltonian, not the derivation of the dipole selection rule breaking down. See section 7.3.) To further optimize the $P$ value, one would like to make the wave functions as long as possible. This is the reason why long organic molecules often can have much larger dipole moments than simple inorganic molecules. As far as “engineering” new electromagnetic materials, however, there is a tradeoff. The smaller dipoles can, in general, be packed more densely into a volume. Perhaps a more fundamental tradeoff would be that increasing $P$ increases the spontaneous rate equally with the stimulated rates. This effect can go either way, as far as the resultant macroscopic effect is considered.

Another interesting salient feature of (9.64) is the factor of $\nu^3$ in the numerator. It was for this reason that, in the mid 1950s after the development of the maser, many thought that there would never be a laser. The spontaneous rate was thought to be too great, and therefore it was generally assumed that it would be too hard to achieve population inversion. As it turns out, laser action is actually easier to achieve than is maser action due to the increased gain possible at optical frequencies as well as the availability of pumps to somewhat efficiently invert the medium. But more on this later. There is also the fact that optical photon energies are much greater than room-temperature phonon energies, allowing for room-temperature laser generation and quantum mechanical detection of optical photons.
Correlation Function for Broadband Emission

An interesting calculation to be carried out in this subsection is that of the field correlation function for the thermal field we are presently considering. The temporal correlation function that was introduced in chapter 3 generally takes the form

\[ R(t_1, t_2) = \langle b(t_1) b^*(t_2) \rangle. \]

In a cavity that is large compared to a wavelength, the spacing in frequency between modes can be quite small compared to the center frequency. As was discussed in an earlier section in this chapter or cold be deduced from the arguments of this chapter concerning the radiation in an empty cavity, the walls of a blackbody cavity reemit the energy that they absorb back into the cavity as a random sequence of photon emissions. The randomness of the emissions will cause each mode to broaden in frequency space as the modal will have a frequency spectrum. For a large enough cavity, the spectral broadenings in each mode will cause the spectra around each eigenfrequency \( \omega_i \) to overlap the spectra surrounding adjacent eigenfrequencies. In such a case, one will see a gain curve, that is, a region in frequency space on which the walls absorb and reemit rather than a set of separate spectral lines. In such a case, it is unnecessary to consider a mode sum, but rather what is effectively a single line. We can then drop the subscript \( i \). We can then write that

\[ R(t_1, t_2) = \langle b(t_1) b^*(t_2) \rangle \] (9.66)

and, using the solution of the equation

\[ \frac{db}{dt} + i\omega b + \kappa b = f(t), \] (9.67)

where

\[ \langle f(t) f(t') \rangle = 2\bar{n}\kappa\delta(t - t'). \] (9.68)

That solution is given by

\[ b(t) = e^{-i\omega t} e^{-\kappa t} \int_t^{t_2} f(t') e^{i\omega t'} e^{\kappa t'} dt'. \] (9.69)

We can then write that

\[ R(t_1, t_2) = \langle b(t_1) b^*(t_2) \rangle = e^{-i\omega(t_1-t_2)} e^{-\kappa(t_1+t_2)} \int_t^{t_1} df(t') e^{i\omega t'} e^{\kappa t'} \int_{t_0}^{t_2} df(t'') e^{-i\omega t''} e^{\kappa t''} \]

\[ = e^{-i\omega(t_1-t_2)} e^{-\kappa(t_1+t_2)} \int_t^{t_1} df(t') e^{i\omega t'} \int_{t_0}^{t_2} df(t'') e^{-i\omega t''} e^{\kappa t''} \langle f(t') f^*(t'') \rangle \]

\[ = \kappa\bar{n} e^{-i\omega(t_1-t_2)} e^{-\kappa(t_1+t_2)} \int_{t_0}^{\min(t_1, t_2)} dt' e^{2\kappa t'} \]

\[ = \bar{n} e^{-i\omega\tau_c} e^{-\kappa|\tau|}, \] (9.5)

where in the last line we took \( t_0 \) to be \(-\infty\) for definiteness. We note that there is an exponential decorrelation of the field with itself. This will lead to a so-called Lorentzian lineshape upon Fourier transforming (9.69) to find spectral intensity. This calculation will be carried out in section 8.2.3.1. This Lorentzian, however, is hardly if at all damped. In the blackbody case, the spectral width can be a major percentage of the central frequency. For most materials (i.e. crystals), the material gain curve will cut off the spectral width long before the Lorentzian damping will. The result is that the lineshape for a device such as an LED will appear to be almost flat, as was discussed in section 7.5 and, in particular, was illustrated in Figure 8.14. Therefore, to define temporal modes for such a source, a good approximation is that which was made in section 7.5 (or as was first discussed in section 3.5) in which the spectrum is approximated as a rectangle function. In this case, the Karhunen-Loeve equations yield either the exact solution of prolate spheroidal wave functions or the quite accurate approximation of the large rectangle being cut up into smaller rectangles, each of a width \( \tau_c \), the coherence time as was discussed in an appendix of chapter 8.
The Lorentzian shape is also the result for a laser high above threshold, as we will see in an earlier section of this chapter. There, though, the linewidth (spectral extent of the line) will be orders of magnitude smaller and will vary in width as the inverse of the output power. Near threshold, it is hard to talk about the concept of lineshape, as in general many modes are competing for the atomic gain, each mode overlapping in physical space but being distinct in frequency space. However, in a semiconductor for example, the broadening is homogeneous, and only one mode can eventually win the mode competition, so there is fierce mode competition in this near-threshold regime.

9.2.3 Phase Noise and the Laser Line

In the preceding blackbody argument, we have essentially ignored the medium and ascribed all material properties to the cavity walls except in the one argument in which relations between the Einstein coefficients were obtained. In this way, attention was focused on what could really be called pure spontaneous emission noise. But spontaneous emission is also the source for laser action as well as the cause of laser phase noise. In the case of the laser, however, it is to possible to ignore the medium as it is the medium which allows for laser action. Here we will need to consider the full set of semiclassical equations where the photon normalized modal amplitude $b_i(t)$ is coupled to each atomic dipole moment’s amplitude coefficient $p_i(t)$ and inversion parameter $d_k(t)$. That is what we presently plan to do.

A coupled, three-variable, essentially nonlinear system such as the system that couples the modal coefficients and the atomic or band related variables, is not going to be easily analytically solved. (By essentially nonlinear, I mean that the nonlinearity is not contained in additive terms which might be considered small to at least set up a perturbation, but here the coupling terms themselves are fully nonlinear—that is, consist of products of the dynamical variables.) This is the reason that problems such as calculating the laser spectrum near threshold are so complex. The laser biased well above threshold, however, offers a set of additional approximations that make it amenable to a potential theory treatment, as was discussed heuristically in section 6.1. The first assumption really is that the field in the cavity is so intense in one single mode $i$ that the $i$ and $k$ subscripts become unimportant as essentially all radiation will be in a single mode (perhaps from high gain or perhaps from high feedback). A second approximation tied to this is that one can assume homogeneous broadening. This approximation cannot be too good as the transition frequency is a function of the subscript $k$ and doing away with the subscript $k$ drops the dependence of the frequency of oscillation on the pump. For a constant pump level, though, which is what we want here for this discussion of linewidth, this approximation may not be so bad when the pump is constant. It is a bad when considering transient dynamics such as the dynamics which lead to laser chirp.

Now, in this single mode, homogeneously damped limit, we can write (approximately) that

\[
\frac{db(t)}{dt} + (i\omega_b + \kappa)b(t) = -ig^* \sum_k p(t) + \frac{d}{dt} \left( \frac{1 + i\omega_p \tau_p}{\tau_p} \right) dp(t) \right) = igb(t) \frac{d}{dt} \left[ \frac{d(t) - d_{eq}}{\tau_d} \right] = 2i(g^* p(t)b^*(t) - gp^*(t)b(t)
\]

where the fluctuating terms have been ignored in the matter equations. Although we have removed the subscript $k$ from the matter equations, the sum over $k$ still has some meaning in the modal coefficient equation despite our above considerations that damping and pumping are approximately independent of $k$ so that $p(t)$ and $d(t)$ can be treated as if they were single transition entities. Later we will sum the matter equations over $k$ in order to transform the $d(t)$ from being a fractional inversion, into being the number of carriers in the conduction band minus the number of carriers in the valence band.

We now make an approximation that the damping constant of the dipole moments is so large (that the time constant with which the dipole moments decay is so short) and that the single mode in which the field is operating is so close to the (roughly one and only) line center of the transitions under consideration, that
(\frac{1}{\tau_p} >> \omega - \omega_p) so that the \( p(t) \) “slaves” to the field \( b(t) \). With this, one can write that

\[
\frac{dp}{dt} + i\omega p(t) = i(\omega_b - \omega_p)p(t) = i\Delta\omega p(t)
\]  

(9.71)

because we have, by assumption,

\[
\Delta\omega << \frac{1}{\tau_p}
\]

(9.72)

and where we have also assumed that any broadening mechanism is homogeneous and weak compared to the strong matter field interaction. The assumption really is that there is at least one mode that is sufficiently close to the line center that each of the dipole moments, due to large damping, cannot relax on its own but instead must adiabatically follow the field. With this approximation, we can solve for \( p(t) \) in the form

\[
p(t) = \frac{id(t)gb(t)}{i\Delta\omega + \frac{1}{\tau_p}}.
\]

(9.73)

With this, we could rewrite our system of homogeneously broadened and pumped equations in the form

\[
\frac{db}{dt} + (i\omega_b + \kappa)b(t) = \sum_k d(t) \frac{|g|^2 b(t)}{i\Delta\omega + \frac{1}{\tau_p}} + f_k(t)
\]

(9.74a)

\[
\frac{dd(t)}{dt} + \frac{d - d_{eq}}{\tau_d} = -\frac{4}{\tau_p} \frac{|g|^2 d(t)|b(t)|^2}{(\Delta\omega)^2 + \left(\frac{1}{\tau_p}\right)^2},
\]

(9.74b)

where we have again ignored the fluctuating source of \( d(t) \) with respect to that of \( b(t) \). If we now take a sum over \( k \) and apply it to the \( d(t) \) equation as well as to the left hand side of the \( b(t) \) equation, duly noting that \( d(t) \) when summed over \( k \) become a variable \( N(t) \) which is just given by

\[N(t) = N_e(t) - N_v(t)\]

where \( N_e(t) \) is the number of electrons in the conduction band and \( N_v(t) \) is the number of electrons in the valence band, then we can write that

\[
\frac{db(t)}{dt} + (i\omega + \kappa)b(t) = \frac{|g|^2 b(t)N(t)\tau_p}{1 + \Delta\omega\tau_p} + f(t)
\]

\[
\frac{dN(t)}{dt} + \frac{N - N_0}{\tau_d} = -\frac{4\tau_p |g|^2 b(t)^2 N(t)}{1 + (\Delta\omega\tau_p)^2}.
\]

The above system of equations and the assorted variants of it are oftentimes called the *semiclassical laser rate equations*. They do take into account certain aspects of carrier dynamics, but also miss some important salient aspects such as the linewidth enhancement factor (Yariv 1995) and gain compression (Hjelme and Mickelson 1989). No theory can take all into account, and many of the interesting effects a that are overlooked by this semiclassical system can be reinserted by first order perturbation theory. In general, the \( 1/\tau_p \) is much larger than \( 1/\tau_d \). For example, in a semiconductor laser, the \( \tau_d \) is on the order of a few nanoseconds, whereas the \( \tau_d \) is on the order of a few femtoseconds. The above equations do break down at or below threshold, if for no other reason than that they are for a single mode. (Far enough below threshold, they break down due to the fact that there is no feedback. The far-below-threshold equations, however, do appear to be single-mode, as the modes mold together into a continuous spectrum which covers the whole gain spectrum.) Including multimodes is generally tricky although above threshold, one can use the multimode rate equations to describe such situations as the ones that lead to mode locking. The fierce mode competition that occurs near threshold, though, probably requires inclusion of polarization, rapidly varying terms, and an inhomogeneous medium due to the spatial dependence of the carrier dynamics.
The Laser as a Particle Moving in a Potential

To get to a picture of a particle moving in a potential (as was qualitatively discussed in chapters 6 and 7), we need to simplify our above semiclassical description still further. Let’s say that we were to ignore the dynamics of \( d(t) \) and therefore \( N(t) \), at least in the sense that we were to say that changes came “adiabatically;” that is, changes in \( b(t) \) did not excite transients in \( d(t) \). (At this point, we will lose the rather noticeable spectral effect of relaxation peaks as well as the stability effect known as relaxation oscillations.) This allows us to ignore the \( \frac{dN(t)}{dt} \) term. A second simplification that is also necessary to make in order to be self consistent with the adiabatic elimination of the dipole moments is to replace the \( \frac{dN(t)}{dt} \) equation by an effective \( N_0 \) and equilibrium inversion. This replacement is due to the smallness of the \( \tau_p \) and the fact that it has already been assumed that transients, excursions of the \( N(t) \) from the \( N_0 \), are small. With this, one can solve the \( N(t) \) equation to obtain

\[
N(t) = N_0 \left[ 1 - \frac{4\tau_p g^2 b(t)^2}{1 + (\Delta \omega)^2} \right].
\]

Substituting this \( N(t) \) back into the relation for \( b(t) \), we obtain,

\[
\frac{db(t)}{dt} + i \omega b(t) + \kappa b(t) = \left[ \frac{g^2 \tau_p N_0}{1 + i \Delta \omega \tau_p} \left[ 1 - \frac{4\tau_p g^2 b(t)^2}{1 + (\Delta \omega)^2} \right] \right] + f(t).
\]

The \( \Delta \omega \) terms in the denominators should give rise to only higher order terms in \( \Delta \omega \) (which was assumed small in order to eliminate \( p(t) \)), so they can be dropped for the present argument about steady state operation. With this, we can write

\[
\frac{db(t)}{dt} + i \omega b(t) = \frac{G(N_0)}{2} b(t) - C(N_0) b(t)^2 b(t) + f(t)
\]

where the gain and saturation terms, that is, the \( G(N_0) \) and the \( C(N_0) \) are given by

\[
G(N_0) = 2|g|^2 \tau_p N_0 - 2\kappa,
\]

\[
C(N_0) = 4\tau_p g^2 N_0.
\]

The behaviors of \( G(N_0) \) and \( C(N_0) \) are sketched in figure 9.15. The value of \( N_0 \) labelled \( N_{th} \) is given by

\[
N_{th} = \frac{\kappa}{|g|^2 \tau_p}
\]

and is the value at which \( G(N_{th}) = 0 \)

that is, the gain threshold value. When \( N_0 \) exceeds \( N_{th} \), there is positive gain.

To complete our picture of the laser as a particle moving in a potential, we use the ansatz that

\[
b(t) = b_0(t) e^{-i \omega t} e^{i \phi(t)}
\]

to obtain

\[
\frac{\partial b_0(t)}{\partial t} = -\frac{\partial V(b_0)}{\partial b_0} + \Re \left\{ \exp \{i \omega t \} e^{i \phi(t)} \right\}
\]

\[
b_0 \frac{\partial \phi(t)}{\partial t} = \Im \left\{ \exp \{i \omega t \} e^{-i \phi(t)} f(t) \right\}
\]

where the \( V(b_0) \) is given by

\[
V(b_0) = \frac{1}{4} \left[ C(N_0) b_0^4 - G(N_0) b_0^2 \right].
\]
The \( f(t) \) has now been multiplied by a complex phase factor. However, as was discussed in regard to the blackbody Fokker-Planck equations earlier in chapter 7, this poses no real problem. Writing

\[
\tilde{f}(t) = e^{i\omega t} e^{-i\phi(t)} f(t)
\]

we note that

\[
\langle \tilde{f}(t) \rangle = 0
\]

\[
\langle \tilde{f}(t) \tilde{f}^*(t) \rangle = 2\kappa m(T)
\]

\[
\langle \tilde{f}_r^2(t) \rangle = \langle \tilde{f}_i^2(t) \rangle = \frac{1}{2} \langle \tilde{f}(t) \tilde{f}^*(t) \rangle
\]

which are just the relations satisfied by the original \( f(t) \). It is hardly even necessary to include the new notation with the tilde as far as statistics are concerned. What we have found is that the \( b_0(t) \) and the \( \phi(t) \) both satisfy Langevin equations. The \( b_0(t) \) equation contains a potential that constrains the motion of \( b_0(t) \). Much as with the blackbody we considered earlier, the motion of the \( \phi(t) \) is undamped, that is, the potential is just a constant. Each time that the Langevin source \( h=g \) gives \( \phi(t) \) a kick, then \( \phi(t) \) picks up a new velocity, which is to say, a new frequency of oscillation. Such is the nature of laser phase noise. Although the amplitude fluctuations are damped, the frequency of oscillation wanders throughout frequency space, limited essentially only by the extent of the gain curve. A one dimensional trace of the potential is sketched for a few values of \( N_0 \) in Figure 9.16. The far-above threshold version, the only one that is really consistent with all of the approximations ont he the derivations sketched in two dimensions in figure 9.17.

The point of the figures is that there is a threshold point at which the system is no longer stable enough to have a zero output value. Above this threshold, the operating point will fall down into the large trough of Figure 9.17 and then roughly carry out the motion of a roulette ball in the trough of a roulette wheel—that is, the phase is not damped, but the amplitude is constrained to stay in the trough. That the phase is not damped is the cause of the linewidth of even the most well-biased, quietest laser. Each spontaneous event will kick the frequency off its former value, even though the amplitude fluctuations get rapidly damped.

A change in the potential shape, as exhibited in Figure 9.16, whereby a single stable point bifurcates continuously into two minima is the type of bifurcation which is characteristic of the potentials associated with an order parameter undergoing a second-order phase transition in a physical system. For example, the poling transition in LiNbO\(_3\) is a second-order phase transition. The quantity corresponding to the \( r \) in the above equation would be the equilibrium position of the lithium atom relative to the halfway point between niobium atoms. Below the Curie temperature, the lithium cannot stay at the halfway point but will spontaneously decide to move one way or the other and do this smoothly with the lowering of the temperature. Due to this analogy, some speak of laser threshold as being an example of a second-order phase transition in a system far from threshold. Unfortunately, the analogy is only mathematical and has nothing to do with an actual laser. The potential equation above was derived under the approximation that the laser was so far...
Figure 9.16: Sketches of the potential $V(\beta)$ for various values of the parameter $d_0$, from $d_0 = 0$ up to above the $d_{th}$ value that was discussed in connection with the rate equation.
Figure 9.17: The potential for the motion of a particle in two dimensions representing the behavior of the complex laser amplitude above threshold.
above threshold that one need only consider a single mode. Near or below threshold, there are a multitude of modes. Far enough below threshold, each mode is so broad in linewidth that the frequencies of all of the modes overlap. In this regime, the modes share excited spontaneously emitted carriers, and the stimulated radiation is small. As threshold is approached, the lines narrow and mode competition takes place. That is, as threshold is approached, the stimulated terms become large, and each mode competes for the excited carriers. Threshold for a given mode, then, becomes a random function of the past history of the complete spontaneously driven system. Each mode will have a different threshold as well, although the first mode to reach threshold may well not be the one to have the highest steady-state gain and may fall again below threshold with increasing bias. Threshold behavior in a laser is not smooth by any means, as the mode competition leads to large fluctuations in output power and wavelength.

The Fokker-Planck Equation for a Laser Well Above Threshold

As we discussed in chapter 7 and again in chapter 8 when conditional Poisson densities were discussed, one can associate a Fokker Planck equation with a Langevin equation so long as the Langevin equations time evolution can be described with a potential function. This is, indeed, the case with the equation for $b_0(t)$. The Fokker-Planck equation for the probability density function (pdf), $p_b(b, t)$, associated with the Langevin equation for $b_0$ (we have dropped the zero from the pdf) is then given by

$$\frac{\partial p_b(b, t)}{\partial t} = \frac{\partial V(b)}{\partial b} p_b(b) + \frac{\kappa m(T)}{2} \frac{\partial p_b(b, t)}{\partial b},$$

where the steady state density satisfies

$$\frac{\partial p_b(b)}{\partial b} = -\frac{2}{\kappa m(T)} \frac{\partial V(b)}{\partial b} p_b(b),$$

which can be integrated to yield

$$p_b(b) = \frac{1}{Z} \exp \left\{ -\frac{2}{\kappa m(T)} V(b) \right\}$$

where the partition function $Z$ is given by

$$Z = \int_0^\infty db \exp \left\{ -\frac{2}{\kappa m(T)} V(b) \right\}.$$

In order to find the conditional Poisson density for the laser well above threshold, we really need to the conditioning density $p_m(m)$ for the number of photons received during a detector period. But this density is just the density for $|b_0(t)|^2$ that can be readily obtained from the density above $p_b(b)$ for the photon amplitude $b_0(t)$. If we make the identifications that

$$A = \frac{|g|^2 \tau_p N_0 - \frac{1}{2}}{\bar{m}(T)},$$

$$B = \frac{|g|^4 \tau_p^2 \tau_d N_0}{\bar{m}(T)},$$

then we can rewrite

$$p_m(m) = \frac{1}{Z} \exp \left\{ A m - B m^2 \right\},$$

$$Z = \int_0^\infty dm \exp \left\{ A m - B m^2 \right\}.$$

The characteristic function for $p_m(m)$ can then be written in the form

$$\psi_m(\omega_m) = \frac{1}{Z} \int_0^\infty dm \exp \left\{ j\omega_m m + A m - B m^2 \right\}.$$
The lower limits on the integrals of equations (9.96b) and (9.97) pose a problem. Completing squares in the integral will lead to error functions which are tabularized—not analytically ascribable. However, equation (9.85) has already assumed that the laser is well above threshold—that is, \[ A \gg B. \] (9.98)

In such a limit, the distribution of \( p_m(m) \), which has a mode of \[ m = \frac{A}{2B} = \frac{A}{A + \frac{1}{2} \frac{1}{2C}}, \] (9.99)
will peak well out from the origin. The value at the origin may well be small, and the relative amount of the distribution for negative values of \( m \) may be negligible. In this limit, then, it should be acceptable to set the lower limits on the integrals of (9.96b) and (9.97) to \(-\infty\). Writing
\[ Z \approx \int_{-\infty}^{\infty} dm \exp \{ Am - Bm^2 \}, \] (9.100)
we can complete the square to find
\[ Z = \exp \left\{ \frac{A^2}{4B} \right\} \int_{-\infty}^{\infty} dm \exp \left\{ -B \left( m - \frac{A}{2B} \right)^2 \right\}, \] (9.101)
where the integral can be evaluated to find
\[ Z = \sqrt{\pi/B} \exp \left\{ \frac{A^2}{4B} \right\}. \] (9.102)

The characteristic function can then be written in the form
\[ \psi(\omega) \approx \frac{1}{Z} \int_{-\infty}^{\infty} dm \exp \{ j\omega m + Bm - Bm^2 \}. \] (9.103)

Completing the square in the integral, we find
\[ \psi_m(\omega) = \frac{1}{Z} \exp \left\{ \frac{(A + j\omega)^2}{4B} \right\} \int_{-\infty}^{\infty} dm \exp \left\{ -B \left( m - \frac{j\omega}{2} + A \right)^2 \right\}, \] (9.104)
which gives the result
\[ \psi_m(\omega) = \exp \left\{ j\omega \frac{A}{2B} \right\} \exp \left\{ -\frac{\omega^2}{4B} \right\}. \] (9.105)

Expanding to second order in \( \omega \), we find
\[ \psi_m(\omega) = \left( 1 + j\omega \frac{A}{2B} + \frac{(j\omega)^2}{2} \frac{A^2}{4B^2} + \cdots \right) \left( 1 + \frac{(j\omega)^2}{2} \frac{1}{2B} + \cdots \right), \] (9.106)
which can be rewritten as
\[ \psi_m(\omega) = 1 + j\omega \frac{A}{2B} + \frac{(j\omega)^2}{2} \left( \frac{A^2}{4B^2} + \frac{1}{2B} \right), \] (9.107)
which gives the following moments, variance, and SNR:
\[ m_1 = \frac{A}{2B} \] (9.108a)
CHAPTER 9. SOURCE AND RECEIVER NOISE

\[ m_2 = \frac{A^2}{4B^2} + \frac{1}{2B} \]  (9.108b)
\[ \sigma^2 = \frac{1}{4B} \]  (9.109)
\[ \text{SNR}_0 = \sqrt{\frac{A^2}{B}}. \]  (9.110)

To find the count distribution \( p_k(k) \) corresponding to the \( p_m(m) \) of the laser distribution, it is easiest to first find \( \psi_k(\omega) \) from the relation

\[ \psi_k(\omega) = \psi_m \left[ -j \left( e^{j\omega} - 1 \right) \right] \]  (9.114)

to obtain

\[ \psi_k(\omega) = \exp \left[ \frac{A}{2B} (e^{j\omega} - 1) \right] \exp \left[ \frac{e^{j\omega} - 1}{4B} \right]. \]  (9.115)

Expanding to second order, we find

\[
\psi_k(\omega) = \left( 1 + \frac{A}{2B} j\omega + \frac{A}{2B} \frac{(j\omega)^2}{2} + \frac{A^2}{4B^2} \frac{(j\omega)^2}{2} \right) \left( 1 + \frac{1}{2B} \frac{(j\omega)^2}{2} \right),
\]  (9.116)

which can be simplified to yield

\[
\psi_k(\omega) = 1 + j\omega \frac{A}{2B} + \frac{(j\omega)^2}{2} \left[ \frac{A}{2B} + \frac{A^2}{4B^2} + \frac{1}{2B} \right],
\]  (9.117)

which leads to

\[ m_1 = \frac{A}{2B} \]  (9.118a)
\[ m_2 = \frac{A^2}{4B^2} + \frac{A}{2B} + \frac{1}{2B} \]  (9.118b)
\[ \sigma^2 = \frac{A}{2B} + \frac{1}{2B} = m_1 + \frac{1}{2B} \]  (9.119)
\[ \text{SNR}_0 = \sqrt{\frac{m_1^2}{m_1 + \frac{1}{2B}}}. \]  (9.120)

The first term in the denominator clearly represents the shot noise contribution. The second term in an excess noise term due to the fact that the conditioning density is not quite a delta function, but instead a density with a smooth minimum. The marble that we have used to represent the state of the intensity is free to roll about the smooth minimum region. Indeed, as a \( A \) and \( B \) are proportional to the \( N_0 \), the pump value of the inversion, we see that the noise term is proportional to the pump level and therefore also to the optical intensity. We will discuss this term further in the section on relative intensity noise.

9.2.4 Spectral Density and Temporal Modes of a Laser Well Above Threshold

Lorentzian Lineshape

The undamped nature of the phase perturbation which was so well illustrated in Figure 8.17 allows us to calculate what the inner lobe of the laser’s line—that is, spectral density—should appear like. (See, for example, Mickelson 1992.) The Wiener-Khintchine theory can be expressed in the form (recall Chapter 3)

\[ S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} R(\tau) e^{i\omega \tau} d\tau, \]  (9.121)

where the correlation function is defined by

\[ R(\tau) = \frac{1}{2\eta_0} \langle E^*(t + \tau) \cdot E(t) \rangle, \]  (9.122)
where the \(1/2\eta_0\) is included to give the correlation functions units of the Poynting vector. Ignoring the amplitude fluctuations, whose actual effect on the spectral density will be to include a quite flat and low-amplitude noise pedestal, we can simplify (9.121) to the form

\[
S(\omega) = \frac{a^2}{2\pi\eta_0} \int \langle e^{i\Delta \phi(\tau)} e^{i(\omega-\bar{\omega})\tau} \rangle d\tau.
\]  

(9.123)

Assuming that the \(\Delta \phi\) is caused by a multiplicity of spontaneous emission events and thereby invoking the central limit theorem, we can express the density of the phase fluctuations \(p(\Delta \phi)\) as a Gaussian:

\[
p(\Delta \phi) = \frac{1}{\sqrt{2\pi}\langle (\Delta \phi)^2 \rangle} \exp \left\{-\frac{(\Delta \phi)^2}{\langle (\Delta \phi)^2 \rangle}\right\},
\]

(9.124)

as we did previously for an incoherent source but for a different reason. As the \(\Delta \phi\) process is the result of a random walk process, we can safely write that

\[
\langle (\Delta \phi)^2 \rangle = 2c|\tau|,
\]

(9.125)

which, together with the use of (9.125), gives

\[
S(\omega) = \frac{a^2}{2\pi\eta} \int e^{-i(\omega-\bar{\omega})\tau} e^{-c|\tau|} d\tau.
\]

(9.127)

The integral can now be performed to yield the Lorentzian lineshape

\[
S(\omega) = \frac{K}{(\omega - \bar{\omega})^2 + c^2},
\]

(9.128)

which is indeed the shape of the central portion of a laser line. This is a lineshape we considered in section 3.5 of Chapter 3 when using the Karhunen-Loeve equations to find temporal modes. There we wrote \(R(\tau)\) in the form

\[
R(\tau) = Pe^{-\Delta \omega \tau}
\]

(9.129)

(i.e. using (9.125) for the \(\langle (\Delta \phi)^2 \rangle\) of (9.124) and suppressing the sinusoid). That is, we used field quantities to derive the \(R\) in section 3.5, and therefore the coefficient becomes a power rather than the count \(\bar{m}\) that we obtain by deriving \(R(\tau)\) with normalized field coefficients. We have suppressed the carrier term, as we were interested there in calculating temporal modes, and we have used \(\Delta \omega\) in place of \(\kappa\) (or \(2c\) in equation (9.100)).

In section 3.5, we found that the eigenmodes and eigenvalues were given by

\[
\varphi_i(t) = \begin{cases} 
\sqrt{2} \left[1 + \sin \left(\frac{b_i t}{\Delta \omega_{\tau_0}}\right)\right]^{1/2} \cos b_i t \\
\tan \frac{\pi b_i}{2} = \frac{\Delta \omega_{\tau_0}}{b_i} \\
\sqrt{2} \left[1 + \sin \left(\frac{b_i t}{\Delta \omega_{\tau_0}}\right)\right]^{1/2} \sin b_i t \\
\tan \frac{\pi b_i}{2} = -\frac{b_i}{\Delta \omega_{\tau_0}}
\end{cases} \quad i \text{ odd}
\]

\[
\varphi_i(t) = \begin{cases} 
\sqrt{2} \left[1 + \sin \left(\frac{b_i t}{\Delta \omega_{\tau_0}}\right)\right]^{1/2} \cos b_i t \\
\tan \frac{\pi b_i}{2} = \frac{\Delta \omega_{\tau_0}}{b_i} \\
\sqrt{2} \left[1 + \sin \left(\frac{b_i t}{\Delta \omega_{\tau_0}}\right)\right]^{1/2} \sin b_i t \\
\tan \frac{\pi b_i}{2} = -\frac{b_i}{\Delta \omega_{\tau_0}}
\end{cases} \quad i \text{ even.}
\]

(9.130)

We further noted at that time that, for narrow lines such that \(\Delta \omega_{\tau_0} < 1\), essentially all of the energy would be detected in the lowest-order mode. For a single mode of a thermal source or an LED, this probably cannot ever be the limit in which we are working, as inverse linewidths of LEDs (ca. 500 fsec) is much shorter than the shortest detector times of ca. 10 psec. But we have already seen that there is another way to get at the temporal modes for thermal sources.
CHAPTER 9. SOURCE AND RECEIVER NOISE

Modes of a Multimode Laser

For a thermal source, a below-threshold source, or simply a multimode source, there will be a number of longitudinal modes in the cavity if there is any nonzero reflection at all from endfaces. We must remember that these longitudinal modes are not necessarily statistically independent and are not the Karhunen-Loeve temporal modes. Karhunen-Loeve modes depend among other things on the detection process. This multimodedness will really have two consequences. One is that, although we keep saying that there are many modes in the cavity, far enough below threshold the $\kappa$'s (the $1/\tau_{bi}$'s) will be sufficiently large that these modes will all spectrally overlap. In this regime also, these modes will share the same atoms for generation. When the modes spectrally overlap, it becomes hard to determine which mode is which, and in this limit the overall lineshape will approach that of the lineshape of any of the single modes. Very far below threshold, of course, the lineshape may well begin to approach the medium gain curve. In this limit, we may as well note that, although the gain curve has some shape, it is quite broad in frequency space, and therefore the temporal coherence time can be significantly smaller than the detector time. This is a limit in which we can pretty much make the Gaussian white noise approximation we discussed in section 7.5 and thereby take the modes as being

$$\varphi_n(t) = \frac{1}{\tau_c} \text{rect} \left[ t - \left( n - \frac{1}{2} \right) \tau_c \right]$$ (9.131a)

$$\lambda_n = \frac{P_{\tau_d}}{\tau_c + 1}.$$ (9.131b)

Finding the temporal modes in any more general case is a complicated problem, but can, in principle, be carried out in the following manner. To illustrate the nature of the problem, let’s consider a spectral density of the form of a sum of Lorentzians, as we depicted schematically in Figure 6.11(b):

$$S(\omega) = \sum_{i=1}^{N} \frac{2\Delta\omega_i P_i}{(\omega - \omega_i)^2 + \Delta\omega^2},$$ (9.132)

where we have included the carriers and where the $P_i$’s must satisfy a constraint such as that the total count $m$ will be given by

$$m = \eta \int dA_d \int d\tau_d \sum \frac{P_i}{\hbar \omega_i},$$ (9.133)

where $A_d$ is the detector area. If we apply the Wiener-Khintchine theorem to $S(\omega)$, we obtain

$$R(\tau) = \sum P_i \frac{\sin \Delta\omega_i (t - t')}{\Delta\omega_i (t - t')} e^{-i\omega_i \tau},$$ (9.189)

where the carriers are nothing to worry about, so one can easily show oneself that each carrier in the $R(\tau)$ will just add that carrier to the associated Karhunen-Loeve mode—at least if there were but one line. That is, if we solved the Karhunen-Loeve equation separately for each line, we would find that

$$\lambda_{ij} \varphi_{ij}(t) = \int P_i \frac{\sin \Delta\omega_i (t - t')}{\Delta\omega_i (t - t')} e^{-i\omega_i (t - t')} \varphi_{ij}(t') dt'$$ (9.190)

would yield us modes of the previously found sort, but now with added carrier. But clearly, these modes will not solve the Karhunen-Loeve equation

$$\Lambda_j \Phi_j(t) = \int \sum_{i=1}^{N} P_i \frac{\sin \Delta\omega_i (t - t')}{\Delta\omega_i (t - t')} e^{-i\omega_i (t - t')} \varphi_j(t') dt'.$$ (9.191)

However, we could assume that the $\Phi_j(t)$ are linearly related to the $\varphi_k(t)$ (we can always find a mapping to take $ij \rightarrow k$) and therefore write

$$\Phi = T \varphi$$ (9.192)
and plug back into the above Karhunen-Loeve equation. As was mentioned previously, the LED line will have a similar shape but with an orders-of-magnitude larger width and for a rather different reason than that the laser has its Lorentzian shape. The result will be an algebraic eigenvalue/eigenvector equation for the rows of the matrix $I$. There seems to be no obvious reason that this procedure would not give the new set of “coordinates” or linear combinations of the lines that move collectively in the statistical sense.

**Mode Partition Noise**

The other effect the multimodedness can have is in the limit where the medium gain is high enough and cavity feedback strong enough that we need to consider the individual modes. This case also arises for multimode lasers. In an inhomogeneously broadened medium—that is, where the modes do not share the same parts of the amplifying medium, different modes can independently exist. Even in a homogeneously broadened medium, mode competition can be so strong that the modes will, so to speak, “flicker.” That is, the lasing mode will change so rapidly that one cannot resolve which is lasing. Therefore the output appears to be multimode. This is a limit in which, although we may well see multiple modes on a spectral trace, there is really only one independent mode. Here there can be an effect known as *mode partition noise*.

Writing the field of a multimode system as a sum of Karhunen-Loeve modes, we find

$$E_x(t) = \sum a_i(t) \phi_i(t)$$

(9.134)

where the $E_x(t)$ is the coefficient of the (assumed single) spatial mode of the field, then we will have something like

$$\int_{t-\tau_d}^{t} \langle E_x(t')E_x^*(t') \rangle dt' = \int_{t-\tau_d}^{t} \sum_{i=1}^{N} \langle |a_i(t')|^2 \rangle dt' = m(t).$$

(9.135)

If the bias to the laser is kept constant, though, it may be that there is only one temporal mode—that is, the individual lines under the gain curve are constrained to have a constant power output. This isn’t to say that the power keeps the same distribution between the lines, though, even in an inhomogeneously broadened medium. This can lead to a time-dependent variation of the spectral content of the laser output. As we will discuss in the next chapter on propagation, this effect can lead to time-dependent dispersion in the transmission channel. The effect can be exacerbated by modulation. However, as will be discussed in the next section, the mode partition effect can often be completely dominated by another dispersive effect known as *chirp*. It should be mentioned here that it is rather hard to predict the magnitude of the partition noise a priori. There are too many possible contributing elements such as the quality of the contacts, ohmic layers, heterojunction interfaces, material purity, unwanted diffusions during processing, geometrical variation, etc., that the effect pretty much has to be characterized experimentally after fabrication.

**9.3 Effects of Modulation**

**9.3.1 Modulation Chirp**

*Modulation chirp* is a rather fundamental effect which relates to the fact that there is always a relationship between the real and imaginary parts of the index of refraction (Koyama and Iga 1988). The semiclassical equations clearly show this, as they contain the complex polarizability $\alpha$. The real part of this variable relates to the real part of the index of refraction, whereas the imaginary part relates to absorption, the imaginary part of the index. In particular, when one injects current into a semiconductor, increasing the optical absorption, this leads to a decrease in the index of refraction. If one injects current into a semiconductor containing a heterojunction, the semiconductor will exhibit gain in the heterojunction (active) region and will thereby exhibit an increase in refractive index. (See, for example, Mickelson 1993.) That this is so follows from the so-called *Kramer’s-Kronig relations*, which state that, for a causal complex function, the real and imaginary parts of that frequency space representation of that function must be related as a Hilbert transform pair. (See Appendix A of Yariv 1997, for example.) Perhaps the easiest way to see what this time-varying index change does to the wave propagating in the medium is to consider a simple example of
plane wave propagation. Say that $\psi(z, t)$ is the complex representation of a wave propagating to the right in a medium of index $n$. One could write that

$$
\psi(z, t) = \psi_0 e^{ikz} e^{-i\omega t},
$$

(9.136)

where it can be shown that $\psi(z, t)$ is a solution of the one-dimensional wave equation

$$
\frac{\partial \psi}{\partial z} = \frac{i}{v} \frac{\partial \psi}{\partial t},
$$

(9.137)

where $v$ is the phase velocity in the medium which is given by

$$
v = \frac{c}{n}.
$$

(9.138)

Although (9.136) may no longer be correct when $n$ becomes a function of time $n(t)$, (9.137) should still be applicable if $n$ varies on a time scale much longer than the optical period. Separating variables in (9.137), we find that

$$
\psi(z, t) = \psi_0 e^{ikz} \exp \left\{ -ikc \int_0^t \frac{dt'}{n(t')} \right\},
$$

(9.139)

where $k$ is the separation constant. Identifying $kc$ with angular frequency $\omega$, we note that an increase in $n$ leads to a decrease in $\omega$. Another way to look at this effect would be to consider the laser cavity. An increase in index in the cavity causes an effective increase in the length of the cavity. If the increase in length is small enough and carried out slowly enough, then the spectrum of the cavity modes should change continuously. However, a slightly longer cavity will have longer wavelength resonances. There is a bit of a problem with this argument, though. If the semiconductor is 300 $\mu$m long and has an index of roughly 3.5, then the roundtrip time $\tau_{rt}$ for a photon in the cavity would be

$$
\tau_{rt} \approx 2 \frac{10^{-3} \text{m}}{3 \times 10^8 \text{m/s}} = 6 \text{ psec}.
$$

(9.140)

However, even if a pulse were to have a rise time of order 6 psec (?) or if the cavity were much longer, there would still be chirp, as shown in the propagation argument above. The photons do not need to have enough time to see the entire cavity in order to be chirped. This is also in agreement with the fact that semiconductor lasers do not have mirrors per se but instead have cleaved facets. The reflection is due to the index difference between the medium with $n_m = 3.5$ and the air with index 1. The intensity reflection $R$ is then given by

$$
R = \left( \frac{n_m - 1}{n_m + 1} \right)^2 \approx 0.35,
$$

(9.141)

which says that the photon lifetime is hardly as much as a full round trip.

Chirp arises anytime there is a change in the relation between $k$ (which can be thought of as a separation constant which happens to denote the propagation constant) and $\omega$. Certainly such a change occurs when the index of refraction $n(\omega)$ is modulated in any manner as it is clear that we can always write that $k(\omega) = \omega n(\omega)/c$. Among other a plethora of other things, current injection, the Stark effect and the electroooptic effect can all serve to modulate the refractive index on a “short” time scale. The chirp effect is often described with the use of a chirp parameter $C$ (Agrawal 1995). In a digital (on/off) modulation, the temporal shape of the pulses radiated from a laser (or LED) can often be described by a super Gaussian

$$
E(z = 0, t) = E_0 \hat{e}_t \exp \left\{ -\frac{1}{2} \left( \frac{t}{t_0} \right)^{2m} \right\},
$$

(9.142)

where $E(z = 0, t)$ is the field radiated from the source, with $E_0$ its amplitude, $\hat{e}_t$ its polarization direction, $t_0$ its initial 1/e intensity point, and $m$ the parameter that describes how much the pulse is flattened from a usual Gaussian shape. ($m = 1$ gives a Gaussian.) To include chirp, one writes

$$
E(z = 0, t) = E_0 \hat{e}_t \exp \left\{ -\frac{1 + iC}{2} \left( \frac{t}{t_0} \right)^{2m} \right\}.
$$

(9.143)
The spectrum of the disturbance $E(z=0,\omega)$ is then given by

$$E(z=0,\omega) = E_0 - \hat{e}_t \left( \frac{2\pi t_0^2}{1 + iC} \right)^{1/2} \exp \left\{ -\frac{\omega^2 t_0^2}{2(1 + iC)} \right\}, \quad (9.144)$$

which has a half width $\Delta\omega$ given by

$$\Delta\omega = \frac{(1 + C^2)^{1/2}}{t_0}. \quad (9.145)$$

The value of $C$ for a semiconductor laser can be significant—as large as 7, but always larger than 1.

One might think that an easy solution to the problem of chirp would be to use an external modulator such as an electroabsorption modulator or perhaps lithium niobate directional coupler or Mach-Zhender modulator. As pointed out by Koyama and Iga (1988), the situation is not so simple. An electroabsorption modulator modulates the absorption and therefore lowers its effective index, shortening the modulator and thereby chirping the phase of the exiting wave. A directional coupler is an interferometric device which works by splitting the incident mode into two modes, an even and an odd mode, and then modulates the phase of the odd mode with respect to the even one before recombining the two at the output. This also must cause chirp, as the phase of the sum of two modes will be something like the average of the phases. However, the phase of the even mode stays constant under modulation while that of the odd changes, causing the output phase to modulate. A solution would be to use a configuration in which the phase of the even mode would be sped up an amount equal to that by which the odd mode is slowed down. This would correspond to modulating only one of the electrodes while dc grounding the other, leading to a factor of two length penalty. In a Mach-Zhender interferometric modulator, the incident channel is split into two independent channels. The ac voltage is usually applied between the two electrodes, one on each channel. In this push-pull configuration, one channel is modulated oppositely to the other. In this case, the chirp of the two channels cancels and the modulation can be close to chirp free. there will be more on modulators in Chapters 12–14.

### 9.3.2 Laser Jitter

Laser jitter is an effect in which there is a variable delay in the time it takes for a laser to turn on from the time that the current pulse is applied. This effect can be the limiting one in the data transmission rate of the system. It can also serve as a serious limit on optical sampling sensitivity (Hjelme and Mickelson 1992).

The effect can be pretty well described by just reconsidering Figure 7.11 of section 7.5. Below threshold, the laser is in a noise-like state. When fully turned off, this state is essentially purely quantum fluctuations. This state was describable as a photon amplitude equation driven by a random source. Well above threshold, the laser will be in a single mode. There are still some dynamics going. A real laser will exhibit relaxation peaks in its spectrum several gigahertz from the center of its line. These peaks are due to the random nature of the emission process—the fact that an emission changes the inversion, which lowers the gain, which damps the photon number to restore the gain. This relaxation oscillation process doesn’t have too much effect at the line center, though. The dynamics of the central line are pretty much determined by a driven Van der Pol-type equation in photon amplitude. That is, the dynamics are governed by an equation of the form of the Langevin equation for $b(t)$ in which the potential term contains a saturation term that is cubic in $b(t)$, or perhaps more correctly, contains a dynamic term of the form $|b(t)|^2 b(t)$. The region near threshold, however, has complicated dynamics indeed. There are multiple modes, which begin to not overlap in frequency space but which are still strongly coupled. The polarizability becomes a dynamical entity despite its short lifetime, so our system of equations must have one photon amplitude per mode. As different modes affect different portions of the medium, using a single inversion parameter and polarizability is probably not correct. Using one per atom may be overkill (and hopeless), but perhaps one could get by with one per mode if the equations were nonlinearly coupled to indicate the competition for gain. For twenty optical modes, this would be a coupled system with a real equation and forty-two complex equations. This is, therefore, a coupled nonlinear system with 23 degrees of freedom. Any system of nonlinear equations with more than three degrees of freedom can exhibit so-called chaotic solutions—that is, solutions which diverge from each other exponentially for infinitesimal changes in the initial conditions. (Note that the coupled system of photon number and inversion parameter seems to have three degrees of freedom, but as we saw in
CHAPTER 9. SOURCE AND RECEIVER NOISE

section 6.1.2, the phase of the photon number actually separates out from its amplitude in the single-mode case.) When one “turns on” the laser from below threshold, the initial condition on the laser’s journey to above threshold is a random one, and this initial condition is on a chaotic system. Clearly, the time taken to traverse this chaotic region will be a random variable. In a semiconductor laser, the jitter can be from one to several picoseconds.

How can one minimize the jitter effect? An obvious way would be to keep the laser at an above-threshold bias. There are drawbacks to this approach. One is power. Even when one isn’t transmitting, one is consuming power. Further, this power is being converted to light in the system. This causes more noise in the system and, for a given system performance, will require the upper state of the laser to have a higher value in order to achieve the same discrimination as with the below-threshold system. An alternate solution would be to run the laser continuous-wave (CW) and use an external modulator with less jitter than the laser. Unfortunately, external modulators tend to be quite expensive and, further, exact a loss penalty in that some amount of light will be lost in coupling into and back out of the modulator.

It should be noted here that it is hard to predict the amount of jitter a priori, at least in a semiconductor. In lasers with external mirrors, the thermal jitter due to changes in length may dominate, or in dye lasers jitter may be dominated by fluctuations in the stream from the dye jet. In a semiconductor laser, however, the gain and amplitude of the oscillation are strongly coupled. The application of a current pulse causes a change in the gain which is slow compared to the optical gain time constants. The laser therefore seems to hesitate before any effect takes place. In a dye laser, the gain can be modulated much faster than the optical time constants (Arrechi et al 1966). In such a case, one can simply look at the initial distribution of states (optical mode amplitudes in the below-threshold case) and calculate the time that a fixed gain will take to breach threshold. In the semiconductor case, the time that threshold is breached is the time that a single mode “wins” the competition of modes in a strongly gain-field coupled system.

9.4 Circuit Noise

Although the formulation of the noise problem in a circuit appears quite different from the blackbody problem, the difference, as we shall soon see, is somewhat illusory. Any time there is a resistance in a circuit, there will necessarily be a corresponding fluctuation, just as the fluctuation dissipation theorem tells us must occurs in an optical cavity. The lower frequency scale posed b the voltage and current oscillations on the electrical circuit does have n effect on the frequency and amplitude distributions of these fluctuations. Also, as the thermal energy $k_B T$ at room temperature corresponds to infrared radiation, one cannot count individual quanta in a circuit. All of these points should become clear in the following discussion.

9.4.1 A Derivation of the Johnson Noise Formula

For the moment, let’s take our model of the photodetector and its circuit to look like that of Figure 9.18(c), which will be sufficient to consider the thermal circuit noise and dark current noise. In general, one considers the load resistor to be the source of the current noise and the back bias on the PiN structure to be the source of the dark current. Although both of these sources are essentially thermal, there is a difference. As we will see in the coming derivation, the circuit noise is due to energy conservation in the sense that the resistor must act like a black body. That is, the electrical energy absorbed must be reradiated into the very short cavity formed between the generating resistor and the load resistor. The absorption and emission are essentially spontaneous processes and are therefore noisy. The noise power therefore becomes dependent only on the temperature of the circuit and the degree of matching between the resistances. The dark current is really due to the detector back bias in the sense that thermal fluctuations can kick an electron up to the conduction band in the detector. This process goes only one way, though, and is therefore not an equilibrium process. The dark current is therefore quite detector-dependent. In this sense, dark current is not as fundamental as the circuit noise that we refer to as Johnson noise.

What we can do to find Johnson’s formula is to consider the circuit in Figure 9.18(c) to be an equivalent cavity resonator—although an admittedly strange one. It might appear schematically as depicted in Figure 9.19. The point is that cavity modes of a “good” resonator would be given by

$$\psi_n^{(\text{res})}(z) = \sin k_n z,$$  (9.146)
with $k_n$ given by

$$k_n = \frac{n\pi}{\ell}$$

(9.147)

or the $n^{th}$ resonant frequency given by

$$f_n = \frac{n v_p}{2\ell},$$

(9.148)

where the velocity of propagation $v_p$ is pretty close to the speed of light. Usually we try to minimize the extent of our circuit for cost as well as radiation considerations. For a 1-cm circuit, the lowest resonant frequency $f_1$ would be roughly 10 GHz. That is to say that we will probably be operating our circuit well below the regime where it appears as any kind of normal resonator. In fact, if we put this 10 GHz into a normal 1-cm circuit, it would radiate away before getting absorbed by the loads. For a 100 GHz signal, one would either want to match the load to the source impedance such a that the signal were absorbed in a single pass, or make the total length $l \ll 1$ cm so that the energy could be absorbed in many passes on the line without any resonance appearing due to the damping time being faster than the inverse resonant frequency. The trick with this circuit resonator is that it’s walls are perfectly black for low enough frequency, and therefore its modes are a continuum below its first resonance. Actually, any resonator into which one can couple in and out will have finite resonance widths in frequency space, as such cavities must be lossy in order to be couplable. Our circuit is so bad a resonator that it has one huge flat resonance and, further, generally contains much less than a single mode.

What is the spectrum of thermal radiation in our circuit? If $R_{\text{eff}}$ is in thermal equilibrium at a temperature $T$ and is in our lossy resonator as well, it must come to equilibrium with any electrical energy in the circuit. The noise that $R_{\text{eff}}$ will be supplying to the load will correspond to the forward-going wave that its
spontaneous emission is supplying to the circuit. If the circuit is effectively filtered to a frequency spectral width $\Delta f$, then the number of modes allowed in only the forward direction is given by

$$\text{# of modes} = \frac{1}{2} \frac{\Delta f}{\text{mode spacing}} = \frac{\ell \Delta f}{v_p}. \quad (9.149)$$

The average energy per mode will be given by the Planck distribution,

$$\langle E \rangle_\nu = \frac{\hbar \nu}{e^{\hbar \nu/kT} - 1} + \frac{\hbar \nu}{2}, \quad (9.150)$$

where $\hbar \nu/2$ comes from the zero point energy and will be important at high frequency. $kT$ at room temperature is roughly .026 eV. One eV corresponds to a frequency of roughly $2.5 \times 10^{14}$ Hz. Therefore, $kT$ at room temperature corresponds to 6 THz. Even if the circuit were cooled to 1 K, this would still correspond to 20 GHz. I think we can safely assume that $\hbar \nu << kT$, and therefore we can rewrite (9.150) in the form

$$\langle E \rangle_\nu = kT, \quad (9.151)$$

and therefore the energy in the cavity is

$$\text{energy} = \frac{\ell}{v_p} kT \Delta f. \quad (9.152)$$

As it takes the energy a time $\ell/v_p$ to reach the load from the $R_{\text{eff}}$, then the power supplied to the load by $R_{\text{eff}}$ is

$$P_{\text{noise}} = kT \Delta f. \quad (9.153)$$

The power supplied by a component to a load is independent of its internal properties and only depends on its temperature and the circuit bandwidth. The voltage and current fluctuations, however, will be dependent on the resistance values. For example, the power dissipated in the load is

$$P_{\text{load}} = \langle I^2 \rangle R_\ell = \frac{\langle V^2 \rangle R_\ell}{(R_{\text{eff}} + R_\ell)^2}. \quad (9.154)$$

Therefore, if the load and effective resistances are matched, then one notes that

$$\frac{\langle V^2 \rangle}{R} = 4kT \Delta f, \quad (9.155)$$

which could as well be written in the form

$$\langle I^2 \rangle = \frac{2kT}{R_\ell \tau_d}, \quad (9.156)$$

where $\tau_d$ is our detector time. This is just the Johnson noise formula, which we set out to derive.

### 9.4.2 Noise Temperature and Noise Figure

The appearance of the temperature in (9.155) allows one to develop the concept of equivalent noise temperature. In the circuit regime, as previously discussed in the derivation of (9.155), thermal noise sources will generally appear to be quite flat, at least at room temperature where the blackbody curves peak at a 10.6-µm wavelength in the infrared (IR). Infrared doesn’t propagate well on transmission lines and, even at liquid nitrogen temperatures of 77 K, the blackbody peak is still at a roughly 45-µm wavelength—still well into the IR. At higher circuit operation frequencies—that is, radio frequency (RF) or into the microwave/millimeter wave regime, one can no longer measure individual voltages and waveforms but only powers and impedances. An incoming waveform can therefore be characterized by an effective impedance and a power. It is sometimes convenient to break that power up into a signal power, $P_s$, and an effective noise power, $P_n$. In the case where the source and load resistance are impedance matched, the thermally induced noise power would be given by the Johnson noise formula. If the match were imperfect, then one could replace the actual noise
temperature by an effective noise temperature. The effective noise power can be used to define an element noise temperature by

\[ T_n = \frac{P_n}{4k\Delta f}. \]  

(9.157)

Of course, an alternate, perhaps more complete, characterization of the signal could be to define the signal-to-noise ratio (SNR) by

\[ \text{SNR} = \frac{P_s}{P_n}. \]  

(9.158)

Oftentimes in RF or microwave engineering, one would like to characterize circuit elements as to their noise properties, as is discussed in Pozar (1990). An amplifier, for example, could be characterized by its effective noise temperature, or by its noise figure which is defined by

\[ F = \frac{P_{so}}{P_{si}} \]  

(9.159)

where the \( i \) denotes input and \( o \) denotes output in (9.159). The noise temperature of an amplifier can be determined by making two amplifier power output measurements, one where the input to the amplifier is a resistor at temperature \( T_1 \) and the other with a resistor at temperature \( T_2 \) at the input. The power measurements would then yield

\[ P_1 = GkT_1\Delta f + GkT_n\Delta f \]  

(9.160a)

\[ P_2 = GkT_2\Delta f + GkT_n\Delta f. \]  

(9.160b)

If \( T_1 > T_2 \), then we could define a greater-than-unity ratio

\[ Y = \frac{P_1}{P_2}. \]  

(9.161)

and determine the \( T_n \) from

\[ T_n = \frac{T_1 - YT_2}{Y - 1}. \]  

(9.162)

Noting that the noise output of a noisy amplifier with an input at temperature \( T \) is

\[ P_{no} = Gk(T + T_n)\Delta f, \]  

(9.163)

we note that the noise figure \( F \) at the output of the noisy gain component is

\[ F = \frac{P_{si}}{GP_{si}} \]  

\[ \frac{Gk(T + T_n)\Delta f}{kT\Delta f} = 1 + \frac{T_n}{T} \]  

(9.164)

or that one can write that

\[ T_n = (F - 1)T. \]  

(9.165)

The above relations are useful, as they can be used to calculate noise figures and noise temperatures of cascades of elements, where each element can be characterized by \( T_{ni}, F_i, \) and \( G_i \). We will leave as an exercise the derivation of the relations

\[ T_{nc} = T_{c1} + \frac{T_{c2}}{G_1} + \frac{T_{c3}}{G_1G_2} + \cdots + \frac{T_{cn}}{\prod_{i=1}^{n-1} G_i} \]  

(9.166a)

\[ F_c = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1G_2} + \cdots + \frac{F_n - 1}{\prod_{i=1}^{n-1} G_i}. \]  

(9.166b)

One last interesting point to finish up this section is to look at a plot of (9.150), the energy per mode supplied by a source at thermal equilibrium at temperature \( T \). Such a plot is given in Figure 9.20. Clearly, spontaneous emission noise is contained in the linearly increasing portion of the plot and circuit noise in the large flat region.
CHAPTER 9. SOURCE AND RECEIVER NOISE

9.4.3 Probability Distributions for Circuit Noise

What will be important in later considerations will be the actual distribution functions for the current in the circuit. Clearly, the Johnson noise is a zero-mean process and could therefore, if considered complicated enough, be represented as the Gaussian function

\[ p_i(i) = \frac{1}{\sqrt{2\pi i^2}} \exp \left\{ -\frac{i^2}{2\langle i^2 \rangle} \right\}. \quad (9.167) \]

The dark current, as we mentioned previously, is due to a different mechanism and emanates from the detector itself in only one direction. A single dark current event will appear essentially identical to the electron emission due to a single photon incident on the detector surface. In fact, for a given dark current level, it would be quite hard to distinguish that dark current from the current generated by a constant but small photon stream. A current realization due to either dark current or signal should then appear as (where more discussion will be given in the next chapter on receiver current)

\[ i(t) = \sum_{i=1}^{k} h_0(t - t_i), \quad (9.168) \]

where \( h(t) \) is an impulse response function due to the detector and attached circuit characteristics, the \( k \) must be chosen from the \( p_k(k) \) for the field incident on the detector, and the \( t_i \) must have some distribution, which may well be Markov. At a fixed temperature, the dark current conditioning number should be a constant \( m_d \). If a field is also incident on the source simultaneously that is clear enough to also be shot noise-limited with conditioning number \( m_c \), then we note that the counts for the process should be additive such that

\[ k = k_c + k_d, \quad (9.169) \]

and therefore the characteristic functions should simply multiply:

\[ \psi_k(\omega) = \psi_{k_c}(\omega)\psi_{k_d}(\omega). \quad (9.170) \]

If both processes are Poisson, then

\[ \psi_{k_c}(\omega) = e^{m_c(e^{\omega} - 1)} \]
\[ \psi_{k_d}(\omega) = e^{m_d(e^{\omega} - 1)}. \quad (9.171a) \]
\[ \psi_{k_c}(\omega) = e^{m_c(e^{\omega} - 1)}. \quad (9.171b) \]
Then
\[ \psi_k(\omega) = e^{m(e^{j\omega} - 1)}, \]  
(9.172)
with
\[ m = m_d + m_c, \]  
(9.173)
and the count distribution \( p_k(k) \) will again be Poisson. A current realization then would be (9.168) with \( k \) distribution according to a Poisson \( p_k(k) \) with \( m \) given by (9.173). We will see in Chapter 11 how to go from an expression like (9.168) to a current distribution \( p_i(i) \).

### 9.5 \( 1/f \) Noise

A form of noise which shows up in practically any system is that of \( 1/f \) noise. The \( 1/f \) noise spectral density appears as in Figure 9.21, at least for low enough frequencies. For example, in a laser diode, the \( 1/f \) spectrum becomes prevalent at tens of KHz (Hjelme et al 1991). In a circuit, the number could be hundreds of Hertz.

What is the cause of this effect, and why is it so ubiquitous? First, perhaps we ponder what it actually means—for example, in the laser. In this case, it means essentially that the center wavelength does not stay fixed but drifts, albeit only slowly. This is somehow not too surprising, as we are pouring current into the laser and heat is being generated along with light. Thermal fluctuations, whatever they are due to, are quite slow to have an effect. Thermal equilibrium is a theoretical construct that takes an infinite time to truly achieve. In most useful systems we can think of, whether optical, electrical, mechanical, etc., we inject energy and hope to get out work. This will practically always lead to some magnitude of thermal fluctuation and, therefore, some amount of drift.

What can be done about \( 1/f \) noise? Basically, one needs to live with it. However, one need not let it affect such things as measurements. One needs to make measurements away from dc. In measuring the power in a laser beam, therefore, one “chops” the beam at some rate above the \( 1/f \) limit and then phase locks the chopped signal to the chopper frequency, effectively making a dc measurement but at an ac frequency. In electrical circuits, one can always do electrical modulation and lock the receiver to the modulation frequency.

### 9.6 Laser Relative Intensity Noise (RIN)

As we illustrated in Figure 7.11, there are basically three spectra that we deal with in optical communications. If we are using a laser biased well above threshold, the spectrum exhibits a narrow Lorentzian (narrow with respect to the full width of the gain region) sitting on a flat pedestal. Near threshold, there appear multiple lines throughout the gain region, each perhaps a thousand times the width of the above-threshold Lorentzian but each still narrow compared to the total gain spectrum, again each resting on a more or less flat pedestal. In the below-threshold case, there is again a single Lorentzian due to the coalescence of the ever broadening
single Lorentzians of the near-threshold case. However, this Lorentzian can become so broad as to fill the gain curve completely. The result is a slightly curving pedestal but of much higher amplitude than that of the pedestals of the cases in 7.11(a) and 7.11(b). As has been mentioned, the line features are really caused by the undamped phase noise within the usual laser cavity. The pedestal, in the ideal case, is the spontaneous emission noise which is damped due to the change in inversion it causes. In a direct detection system, the receiver does not see phase noise but only intensity noise. Indeed, the Poisson distribution is the distribution of intensity noise. But the noise that is inherent in the Poisson density is proportional to the square root of the number of photons received during the detection period. We generally refer to this noise as counting noise or shot noise. When our conditioning density \( p_m(m) \) in the expression for the conditional Poisson density

\[
p_k(k) = \int_0^\infty dm \, P_{\text{os}}(k, m) p_m(m)
\]

is a delta function then we have only shot noise. We saw earlier that even a laser biased well above threshold has a \( p_m(m) \) that may have a reasonably sharp peak but it is not a delta function. This leads to an excess noise term that we refer to as relative intensity noise or RIN. This noise really has two sources. We saw earlier that this RIN is proportional to the pump \( N_0 \) as is the signal intensity. One source of RIN is that the laser intensity is not fixed to the pump by the potential, but rolls around it. The other source of RIN is the circuit supplying the inversion, as this circuit does not completely fix the \( N_0 \). We will consider both of these effects in coming paragraphs.

Earlier, we had derived equations coupling the modal amplitude \( b(t) \) and \( t \) the (dynamical) degree of inversion. We repeat these equations here in the form

\[
\frac{db(t)}{dt} + i\omega b(t) + kb(t) = \frac{G}{2} b(t) + f(t)
\]

\[
\frac{dN(t)}{dt} + \frac{N(t) - N_0}{\tau_d} = -2G|b(t)|^2
\]

where

\[
G \approx 2|g|^2 \tau_p N(t).
\]

The equations can be rewritten as a system with \( m(t) = b * (t)b(t) \) and \( N(t) \) as

\[
\frac{dm(t)}{dt} = \left( G - \frac{1}{\tau_b} \right) m(t) + F_m(t)
\]

\[
\frac{dN(t)}{dt} + \frac{N(t) - N_0}{\tau_d} = -2Gm(t)
\]

where \( F_m(t) \) is the Langevin source for the intensity fluctuation. These equations are still nonlinear due to the appearance of \( N(t) \) in the \( G \) term. The equations can be linearized by writing that

\[
m(t) = m_0 + \delta m(t)
\]

\[
N(t) = N_0 + \delta N(t)
\]

where the \( \delta m(t) \) and \( \delta N(t) \) can be considered to be much smaller than the mean values \( m_0 \) and \( N_0 \), respectively. The resulting linearized system of equations is given by

\[
\frac{d\delta m(t)}{dt} = \left( G - \frac{1}{\tau_p} \right) \delta m(t) + \delta G(t)m_0 + F_m(t)
\]

\[
\frac{\delta N(t)}{\tau_d} + \frac{\delta N(t)}{\tau_d} = -2G(t)\delta m(t) - 2\delta G(t)m_0.
\]

Clearly, the system can be solved by differentiating the second equation, substituting for \( \frac{d\delta m(t)}{dt} \), and then solving the resulting second order equation for \( \delta N(t) \) by a Green’s function technique, and then plugging back into the \( \frac{d\delta m(t)}{dt} \) equation. Although the technique is straightforward, the details are tedious (Agrawal, 1992). The basicpoint here is that one can find an autocorrelation function

\[
R_{m,m}(\tau) = \langle \delta m(t) \delta m(t + \tau) \rangle / m_0^2
\]
whose Fourier transform gives us the intensity spectral density

\[ RIN(\omega) = \int_{-\infty}^{\infty} d\tau R_m m(\tau) e^{-i\omega\tau}. \]

More information about this equation can be gleaned from the taking the derivative of the \( \frac{d\delta N(t)}{dt} \) equation to obtain

\[ \frac{d^2 \delta N(t)}{dt^2} + \frac{\delta N(t)}{\tau_{RO}} + \omega_{RO}^2 \delta N(t) = 0 \]

where we have ignored the Langevin source. The information that we have gleaned about the spectral function are the values for the relaxation oscillation frequency \( f_{RO} = \frac{\omega_{RO}}{2\pi} \) and the damping constant for these oscillations, the \( \tau_{RO} \). Evidently, the RIN spectrum will have a maximum at this relaxation frequency and will fall off in either direction in frequency space from this peak at a rate given by the damping constant. Typical relaxation oscillation peaks of laser diodes are in the range of several gigahertz. When communications are taking place at modulation rates well below this range, one often takes the RIN to be constant at a typical value of \( 10^{-16} \) (160dB) (Yariv 1997). In this range, one can take the RIN current to be roughly given by

\[ i_{RIN}^2 = e^2 RIN m_0^2 \Delta f. \]

Now, as we had mentioned above, fluctuation of the pump can also generate additional RIN. We will now give some consideration to that effect. The two driving terms in \( b(t), N(t) \) system were a field fluctuating force \( f(t) \) and the inversion \( N_0 \). If the sole source of the pump is a steady current \( I_0 \), then one could write that

\[ N_0 = \frac{I_0 \tau_d}{e}, \quad (9.174) \]

where \( e \) is the electron charge. (9.174), however, cannot be the whole story. As we have seen in the previous section, even an ideal source generates thermal noise due to resistance. The fundamental noise limit of the laser is therefore more than spontaneous emission noise, as the driver is impressing a Gaussianly distributed current on the junction. In general, though, the situation is still worse, as even a steady current source will generate excess noise, as will time-varying parasitics, etc., which in great degree may be due to packaging. For present analysis, a possibility is to lump all of these sources together into an excess ac current noise source \( i(t) \) which adds to the dc source, \( I_0 \).

For a given dc bias, a small signal ac equivalent circuit for a laser may be given as in Figure 9.22, where the source is considered to be just our noise source \( i(t) \) with a short resistance \( R_S \) to ground. The current which actually pumps the inversion we will call \( i_p(t) \), which is the current flowing through the resistive part of the laser junction impedance. One can write a node equation for this ac equivalent circuit to find that

\[ i(t) = \frac{v_L}{R_S} + C_L \frac{\partial v_L}{\partial t} + v_L \frac{R_L}{R_L} + \frac{1}{L} \int v_L(t) \, dt \quad (9.175a) \]

\[ i_p(t) = v_L(t)/R_L, \quad (9.175b) \]

where it should be noted that the laser relaxation oscillation frequency is given by

\[ \omega_R = \frac{1}{\sqrt{L_C}}. \quad (9.176) \]

A Laplace transform of (9.175a) gives

\[ v_L(s) = \frac{si(s)/C_L}{s^2 + \left(\frac{C_L}{R_S} + \frac{C_L}{R_L}\right) s + \omega_R^2}. \quad (9.177) \]

Noting that

\[ \frac{s}{s^2 + \left(\frac{C_L}{R_S} + \frac{C_L}{R_L}\right) s + \omega_R^2} = \frac{A}{s + \alpha + j\omega} + \frac{B}{s + \alpha - j\omega}, \quad (9.178) \]
where $\alpha$ and $\omega$ come from taking the roots of the second-order polynomial, we see that

$$L^{-1} \left[ \frac{s^2 + \left( \frac{C_L}{R_S} + \frac{C_L}{R_L} \right) s + \omega^2}{s^2 + \left( \frac{C_L}{R_S} + \frac{C_L}{R_L} \right) s + \omega^2} \right] = e^{-\alpha t} \cos \omega t,$$  

(9.179)

which in turn gives that

$$v_L(t) = \frac{1}{C_L} \int e^{-\alpha(t-t')} \cos \omega(t-t')i(t') \, dt',$$

(9.180)

and therefore the noise pump current $i_p$ is given by

$$i_p(t) = \frac{v_L}{R_L},$$

(9.181)

or

$$i_p(t) = \frac{1}{\tau_L} \int e^{-\alpha(t-t')} \cos \omega(t-t')i(t') \, dt'.$$

(9.182)

If one then makes the assumption that $i(t)$ is delta-correlated—that is,

$$\langle i(t)i(t') \rangle = e^{2R} \delta(t-t')$$

(9.183)

where $R$ is a noise emission rate, then one can write

$$\langle i_p^2(t) \rangle = \frac{1}{\tau_L^2} e^{-2\alpha t} \int_{-\infty}^{t} e^{\alpha(t+t'')} \cos \omega(t-t') \cos \omega(t' - t'') \langle i(t')i(t'') \rangle \, dt' \, dt'',$$

(9.184)

which can be simplified from (9.183) to yield

$$\langle i_p^2(t) \rangle = \frac{e^{2R}}{2\tau_L} e^{2\alpha t} \int_{-\infty}^{t} e^{2\alpha t'} \, dt$$

$$= \frac{e^{2R}}{4\tau_L} \frac{1}{\alpha \tau_L},$$

(9.6)

which gives the average amount of excess noise coupling into the laser junction.

---

**Figure 9.22**: Laser ac equivalent circuit being pumped by an ac noise source.
Threshold would imply that \( N_0 > \kappa / G \tau_\alpha \) (9.187)

and therefore that

\[
b(t) = \int e^{A(t-t')}e^{-i\omega(t-t')}F(t') \, dt',
\]

where the \( \delta N(t) \) noise term is now in the positive \( A \). Whatever the initial distribution was, it gets transformed by an exponential.

**Problems**

1. An infinite lens (almost) performs a two dimensional Fourier transform of the shape of a light distribution placed a distance \( f \) the lens focal length, in front of the lens, where the locations of the transform pattern is at a distance \( f \) behind the lens. A curved mirror can perform a similar operation on a distribution of light at distance \( f \) in front of the mirror at the same plane. A confocal resonator then perform (approximately again) a Fourier transform operation each pass around the resonator. The stationary light distributions possible in the resonator then must be eigenfunctions of the Fourier transform. Show that the eigenfunctions of the Fourier transform operation are the Gaussian Hermite modes, that is, show that the Gaussian Hermite modes are defined by the integral relation

\[
\psi(x) = \lambda \int_{-\infty}^{\infty} dx' e^{-ikx'} \psi(x').
\]

What are the eigenvalues?

2. In statistical thermodynamics, the partition function

\[
Z = \sum_{n=0}^{\infty} e^{-\mathcal{E}_n / kT},
\]

where \( \mathcal{E}_n \) are the energy eigenvalues, is an important quantity, as it contains all of the knowledge of the spectrum of eigenvalues. Show that

\[
\langle \mathcal{E} \rangle = \frac{\partial Z}{\partial T}.
\]

3. One often writes two-level rate equations in the form

\[
\frac{\partial N_2}{\partial t} = AN_2 + BN_1
\]

\[
\frac{\partial N_1}{\partial t} = CN_2 + DN_1.
\]

(a) What conditions must be satisfied by \( A, B, C, \) and \( D \) in order that \( N = N_1 + N_2 \) is a constant in time?

(b) What steady-state conditions must be satisfied by \( A \) and \( B \) in order that the thermal equilibrium condition

\[
N_2/N_1 = e^{-\hbar\omega / kT}
\]

be satisfied?

(c) Given the initial conditions

\[
N_1(0) = N_{10}, \quad N_2(0) = N_{20},
\]

solve the system of equations.

(d) Can one obtain inversion, i.e. \( N_2 > N_1 \), by simply raising the temperature \( T \)?
4. Use equation (9.33) in Schrödinger’s equation (equation (9.46)) to obtain the relation (9.34) and therefore obtain equation (9.35) for the Einstein $B$ coefficient.

5. Recall that the power received by a diffraction-limited receiver from a blackbody radiator at temperature $T$ is given by

$$P_{b0}(f) = \frac{\lambda^2 \rho(\nu)}{2\tau_d},$$

where the radiance function $\rho(\nu)$ is given by the blackbody formula,

$$\rho(\nu) = \frac{4 h \nu^3}{c^3} \frac{1}{\exp\{h \nu / kT\} - 1},$$

with $\lambda$ the center wavelength, $\tau_d$ the detector time, $\nu$ the center frequency, $h$ Planck’s constant, $c$ the speed of light, and $k$ Boltzmann’s constant.

(a) Sketch $P_{b0}(\nu)$ as a function of $\nu$.

(b) Find the center frequency $\nu$ about which $P_{b0}(\nu)$ is a maximum. Interpret your result.

6. Solve the second-order system of equations

$$\frac{\partial f_1}{\partial t} + \alpha_1 f_1 = f_2(t)$$

$$\frac{\partial f_2}{\partial t} + \alpha_2 f_2 = Qe^{-at}.$$

7. Use the Wiener-Khintchine theorem,

$$S(\omega) = \int e^{-i \omega \tau} R(\tau) d\tau$$

where

$$R(t_1, t_2) = \langle b^*(t_1)b(t_2) \rangle$$

where

$$\frac{\partial b(t)}{\partial t} + \alpha b(t) = f(t)$$

where

$$\langle f(t)f(t') \rangle = \bar{n} \delta(t-t'),$$

to find the $S(\omega)$ for the above-defined process.

8. Solve for a specific realization of the process

$$\frac{\partial b(t)}{\partial t} + i \omega b(t) + \kappa b(t) = f(t)$$

where $f(t)$ may be given by

$$f(t) = \sqrt{\frac{\hbar c}{\tau_d}} \sum \delta(t-t_i).$$

9. The system

$$\frac{\partial b}{\partial t} + i \omega b + \delta b = \frac{g^2 db}{i \Delta \omega + \tau_N}$$

$$\frac{\partial d}{\partial t} - \frac{d-d_0}{\tau_N} = -\frac{4g^2 d|b|^2}{\tau_\alpha \left[ (\Delta \omega)^2 + \left( \frac{1}{\tau_\alpha} \right)^2 \right]}$$

is not generally solvable. Try to solve it in the following limit:
(a) (i) Take $\Delta \omega = 0$; that is, consider the line center.
(b) (ii) Assume that carrier dynamics are unimportant, i.e. take $\partial d/\partial t = 0$. These approximations should yield a first-order nonlinear differential equation. The additional assumption
(c) (iii) $|b|^2 << 1$ should allow one to write down an analytical solution to the problem.

10. The system
\[
\frac{\partial b}{\partial t} + i\omega b + \delta b = \frac{g^2 db}{i \Delta \omega + \tau_\alpha}
\]
\[
\frac{\partial d}{\partial t} + \frac{d - d_0}{\tau_N} = -\frac{4g^2 |b|^2}{\tau_\alpha [(\Delta \omega)^2 + (\frac{1}{\tau_\alpha})^2]}
\]
carries full information about turn-on transients. To see this in the simplest way, first eliminate the nonlinear terms through making an expansion
\[
d = d_{th} + \delta d
\]
\[
b = b_{th} + \delta b
\]
where the $d_{th}$ and $b_{th}$ are the thermal populations and the $\delta d$ and $\delta b$ are small oscillating terms due to a turn on/off $d$ at time $t = 0$ from $d_{th}$ to $d_0$.

(a) Obtain coupled linear differential equations for $\delta b$ and $\delta d$.
(b) Obtain the uncoupled second-order differential equations for $\delta d$ and $\delta b$.
(c) Solve the equations of (b).
(d) What do the new composite damping and frequencies of uncoupled systems mean?

11. The solution of the Fokker-Planck equation for the distribution $p_m(m)$ for a potential $V(m)$ is given by
\[
p_{m}(m) = \frac{1}{Z} e^{-V(m)}.
\]
Recalling that the characteristic function of a conditional Poisson process is (equation (6.82))
\[
\psi_k(\omega) = \psi_m (j(1 - e^{i\omega})) ,
\]
then plot the first and second moments of the conditional Poisson process where
\[
V(m) = \frac{m}{m^2} + \frac{2 m^2}{m^2}.
\]

12. Pick the $\Delta \omega$ and $\tau_N$ in the following,
\[
\phi_i(t) = \begin{cases} 
\frac{\sqrt{2}}{\sqrt{\pi}} (1 + \frac{m b_i \tau_\alpha}{\sqrt{\tau_\alpha}})^{1/2} \cos \frac{b_i \tau_\alpha}{2} & i \text{ odd} \\
\frac{\sqrt{2}}{\sqrt{\pi}} (1 + \frac{m b_i \tau_\alpha}{\sqrt{\tau_\alpha}}) \sin \frac{b_i t}{\tau_\alpha} & i \text{ even}
\end{cases}
\]
such that there are roughly ten modes and then plot these modes out. How do they compare with the modes of an incoherent source?

13. Assume in (9.112–lost) that there are several (perhaps ten) Lorentzian modes under a gain curve. Perform the transformation of equation (9.115–lost) and thereby determine what the statistically independent modes are.
14. Let’s try to solve the system

\[
\frac{\partial \alpha_\mu}{\partial t} + i\omega \alpha_\mu + \frac{\alpha_\mu}{\tau_\alpha} = ig_\mu b_{\text{ext}} d_\mu
\]

\[
\frac{\partial d_\mu}{\partial t} + \frac{d_\mu - d_0}{\tau_d} = 2i(g^*_\mu b^*_{\text{ext}} \alpha_\mu - g_\mu b_{\text{ext}} \alpha^*_\mu)
\]

including transients.

(a) Solve the above assuming

\[
\alpha_\mu = i\tau_\alpha g_\mu d_0 b_{\text{ext}} \approx -i\tau_\alpha g_\mu b_{\text{ext}}
\]

but not ignoring the time derivative of \(d - d_0\) (although one can use the fact that \(d_0\) has a zero time derivative).

(b) Try solving by integrating the \(\alpha_\mu\) equation first and substituting into the \(\frac{\partial}{\partial t}(d_\mu - d_0)\) equation.

15. A typical noise model for a circuit might appear as below, where \(R_1 = R_2\) assumes that the loads are matched so that modes of the line do not reflect back and forth.

16. Generally, one can consider the circuit to be a lossy cavity, and therefore all frequencies down to DC can exist in the cavity (circuit). The average energy in a cavity mode is given by

\[
\langle E_\omega \rangle = hf \left[ \frac{1}{2} + \frac{1}{e^{hf/kT} - 1} \right],
\]

which can be converted to spectral current density simply by multiplying by a constant.

(a) Sketch the dependence of \(\langle E_\omega \rangle\) on \(hf\), paying attention to the regimes where \(hf \ll kT\) and \(hf \approx kT\).

(b) Generally, one assumes \(hf \ll kT\) and that the modes where \(hf \rightarrow kT\) are radiated away from the line. Let’s say, though, that we cryogenically cool our detector circuit. For a direct detect system, we often write that

\[
S_{\text{noise}}(\omega) = \alpha P + \frac{N_t}{e^2},
\]

where \(N_t\) is flat. This leads to an \(\text{SNR}\big|_p\) of

\[
\text{SNR}\big|_p = \frac{\alpha^2 P^2 \chi/2B}{\alpha P + \frac{N_t}{e^2}}
\]

if one assumes a flat bandwidth filter of width \(2B\).

(c) Find an expression for the \(\text{SNR}\big|_p\) when \(kT \ll hf\); that is, for the “interesting” noise modes. When is cooling a useful strategy?

17. Johnson noise is the spontaneous thermal fluctuations in the voltage across a resistor. The mean-squared noise voltage is proportional to the value of the resistance \(R\). Use the one-dimensional Planck law to derive the Nyquist theorem

\[
\langle V^2 \rangle = 4kTR\Delta f,
\]
where $R$ is the resistance of the resistor, $V$ is the voltage across it, $k$ is Boltzmann’s constant, $T$ is the thermal equilibrium temperature of the resistor, and $\Delta f$ is the frequency bandwidth within which the voltage fluctuates. The circuitry following the photodetector can be modeled as in the figure below. Each electron forming in the photodetector contributes a voltage value $GeR/\tau$ to the output of a $\tau$-sec integrator (counter). If $n$ electrons are generated in the $\tau$-sec counting interval, the integrator intensity (voltage squared) is

$$I = n \frac{GeR}{\tau}.$$ 

Here $G$ is the photomultiplier in the detector, and Poisson statistics hold for $I$. Prove that, when one includes the Johnson noise, the probability density of $I$ is given by

$$P_I(I, m\tau) = \sum_{j=0}^{\infty} \frac{m^j}{j!} \exp(-m\tau) G\left(I; j \frac{GeR}{\tau}, \frac{4KT\tau}{\tau}\right),$$

where $G(I : p, q)$ denotes a Gaussian density in the variable $I$ with mean $p$ and variance $q$.

18. Equations (9.134) have some rather interesting consequences. Systems such as the front end of an optical receiver or a radioastronomical antenna will have an amplifier right at the detector (antenna element). What characteristics would one want this amplifier to have according to (9.134)? Interestingly enough, not all systems want to be low noise. The front end of a radar may appear as below.

Military radars generally need very large dynamic range—that is, they’d like to pick up weak signals but don’t want to be knocked out by jammers. Why might a military radar not want an element amplifier at all?

19. Say that the circuit noise in Figure 9.22 is Gaussianly distributed. Use the standard distribution function transformation law to find out what distribution function the noise radiation caused by this circuit noise has.


[17] H. Kroemer,


