

Let us now return to the more general case that $\omega_A \neq \omega_B$. Specifically, suppose that one frequency component represents the frequency, ω_0 , of the spectral peak of interest, and that the second component at frequency, ω_{ref} , is generated by a reference oscillator, ω_{ref} , is chosen sufficiently close to ω_0 that

$$|\omega_{ref} - \omega_0| \ll |\omega_{ref} + \omega_0| \tag{4.15}$$

Next, let the two signals be added to form the input to a suitable non-linear device. The device output (as we have previously noted) will contain signals at frequencies, ω_0 , ω_{ref} , $2\omega_0$, $2\omega_{ref}$, $(\omega_0 + \omega_{ref})$, and $(\omega_0 - \omega_{ref})$, as shown in Figure 4.22. Finally, let the output signal be subjected to a low-pass filter (i.e., a device which will pass signals only up to an upper frequency limit). The low-pass filter removes all of the fundamental, harmonic, and sum-frequency components, and leaves only the much lower-frequency difference-frequency components.

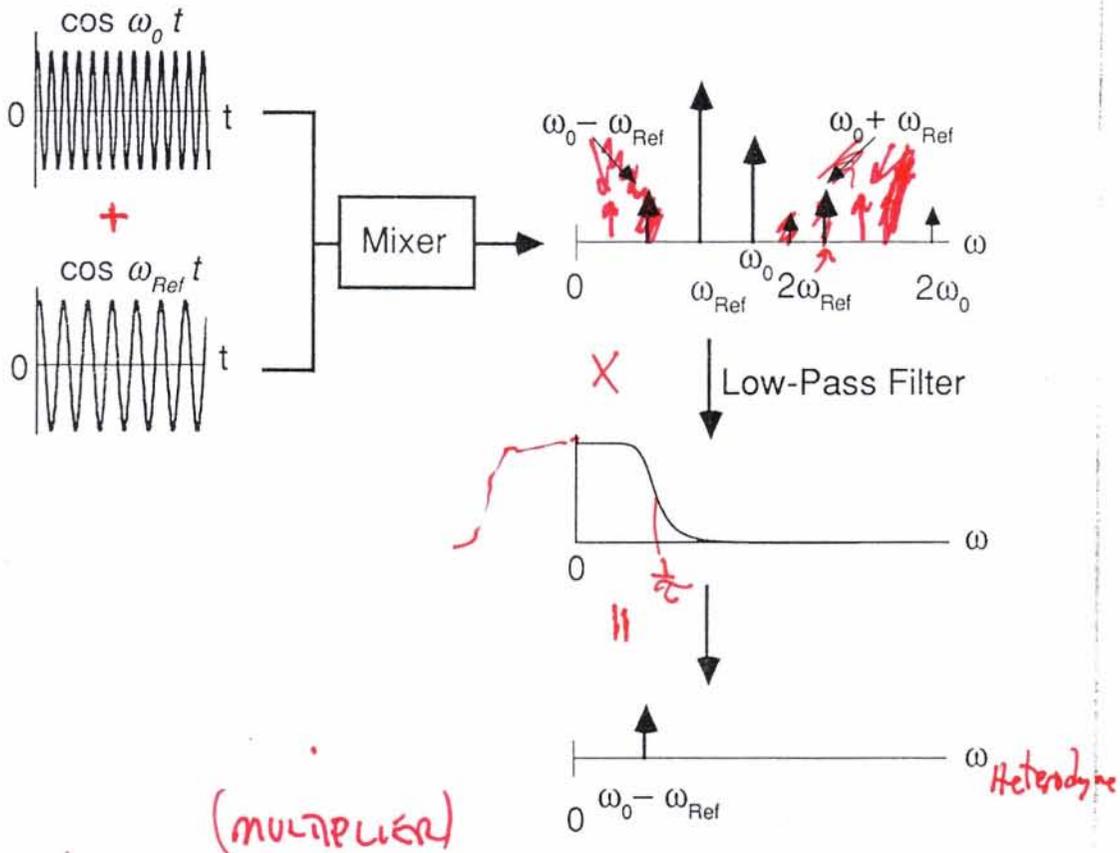
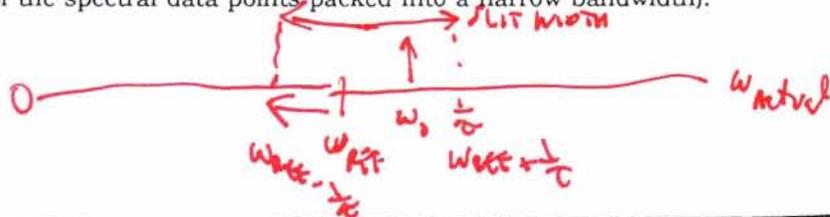
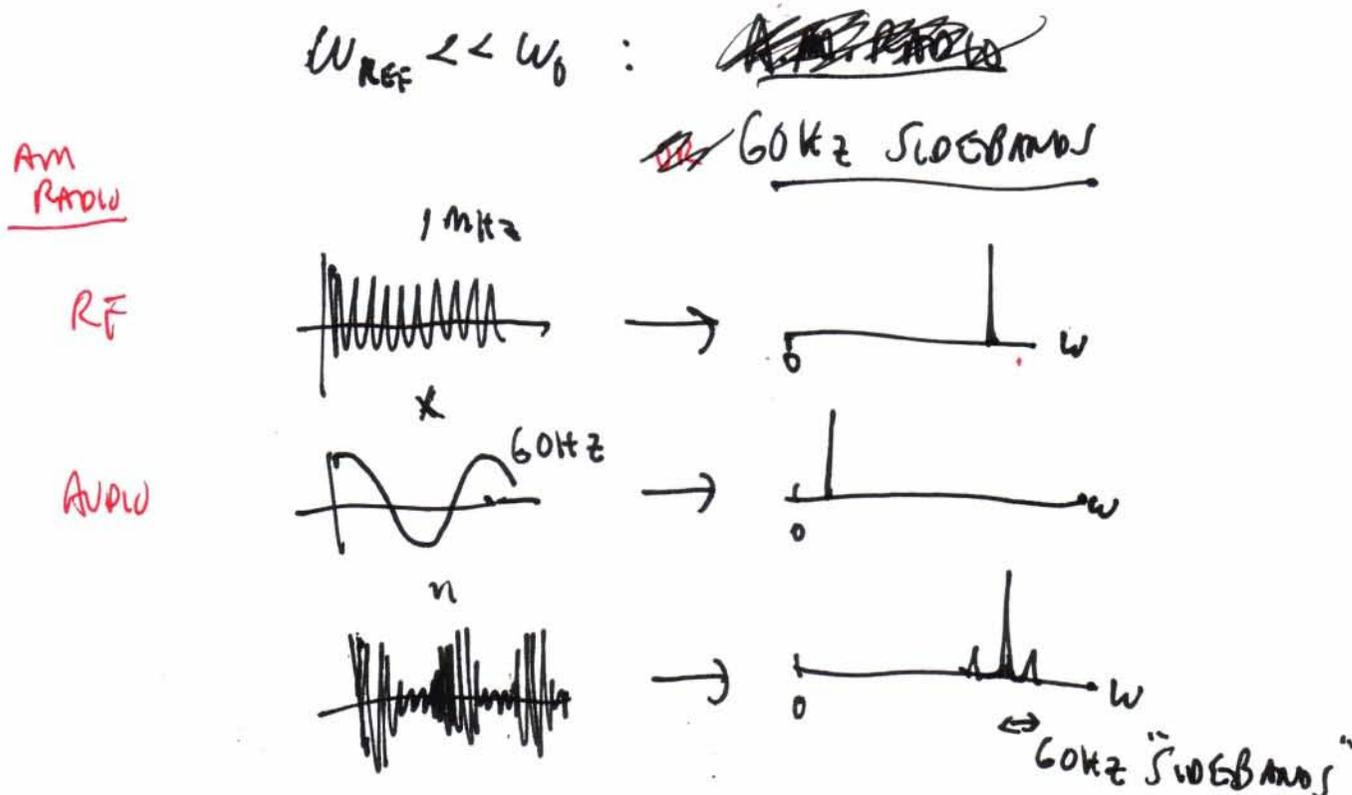


Figure 4.22 Schematic heterodyne detector, consisting of a reference oscillator, mixer, and low-pass filter. The mixer is a non-linear device which has an effect equivalent to multiplying the two input signals together. The mixer output has frequency components at multiple (harmonic) and combination (intermodulation) frequencies. The low-pass filter removes all but the difference frequencies, leaving a narrow-band frequency-domain spectrum of potentially high digital resolution (i.e., all of the spectral data points packed into a narrow bandwidth).



In physical terms, the effect of the mixer/filter combination is to serve as a spectral "slit", whose position is centered at the frequency of the reference oscillator, and whose width is equal to the bandwidth of the low-pass filter. Because the mixer has an effect similar to multiplying ("beating") the two input signals together, the experiment is called "heterodyning".

Apart from any FT applications, the heterodyne experiment has several major advantages for spectroscopy. First, the "slit" width is adjusted electronically, and may therefore be made arbitrarily wide or narrow without any mechanical adjustment to the spectrometer. Second, the reference oscillator frequency may be specified with high accuracy (e.g., a laser for the case of optical mixing, or an rf oscillator for NMR or ICR mixers). Since the frequency scale is also fixed electronically, without internal calibration, it can therefore be extremely accurate.



Although computation of the C -value in Eq. 6.64 affords a means for comparing trial and experimental data, the C -criterion alone is (perhaps surprisingly) not sufficient to establish the "best" spectrum. For example, one might at first intuitively infer that the spectrum produced by direct FT of the actual time-domain data should be the "best" spectrum, because the inverse FT of that spectrum yields the original time-domain data set *exactly* (i.e., $C = 0$ in Eq. 6.64). In the remainder of this chapter, we shall try to expand our intuition to show that we need a second criterion which will turn out to be maximization of the entropy of the trial spectrum.

6.4.2 Probability peaking: most probable result

Faced with uncertainty in choice of the "best" spectrum, we should not be surprised to find that the "best" answer is a matter of probability. In general, the probability, $P(m)$ of a particular result, m , is defined by:

$$P(m) = \frac{\text{Number of ways of obtaining the result } m}{\text{Total number of ways of obtaining all possible results}} \quad (6.66)$$

For example, in one throw of two (honest) dice, there is only one way to roll a 2 (namely, 1,1), two ways to roll a 3 (1,2; 2,1), and so on. From Eq. 6.66, the reader can easily show that the odds of rolling a 7 are thus, $6/(1+2+3+4+5+6+5+4+3+2+1) = 6/36 = 1/6$. The number, 7, is also the most probable result, because there are more ways (6) of obtaining it than of any other number between 2 and 12.

Example: Coin tosses

~~An MEM spectrum (see below) is the most probable spectrum which matches the experimental time domain data to within some previously specified criterion (e.g., a particular maximum C -value in Eq. 6.64).~~ However, we first need to show that only the *most probable* result matters, so that we can ignore all other possible outcomes. Therefore, consider first the simpler problem illustrated in Figure 6.13. The relative probability of obtaining m heads after N tosses of a single coin is given by:

$$P_N(m) = \frac{N!}{m!(N-m)!} \left(\frac{1}{2}\right)^N \quad (6.67)$$

When the number of tosses is relatively small (say, $N < 10$), the number of heads from a particular experiment can differ significantly from the most probable result, namely, $N/2 = 5$ heads. However, as the number of tosses becomes large, we will be very likely to observe a result very close (percentage-wise) to the most probable result (i.e., $N/2$ heads in a particular N -toss experiment). Thus, for large N , we can safely ignore all but the *most probable* result.

The result shown in Figure 6.13 is a direct consequence of the *number of ways* of producing a given result (in this case, the relative number of heads after N coin tosses): e.g., there is only one way that *all* of the tosses can be heads, but there is a huge number of ways that *half* of the tosses can be heads. For large N , we are therefore almost certain to obtain $\sim N/2$ heads.

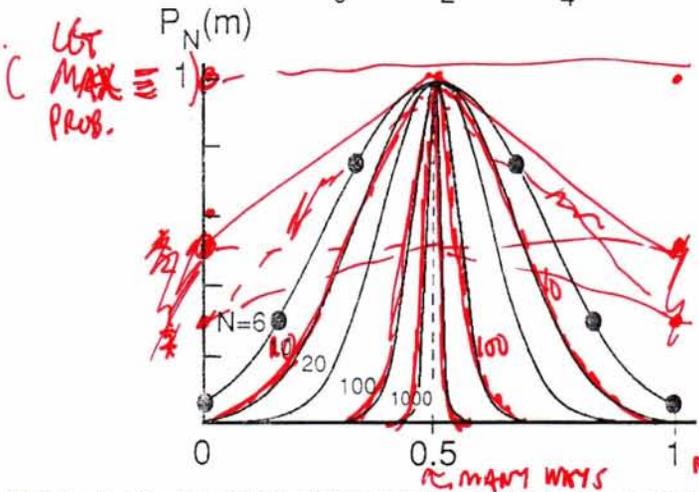
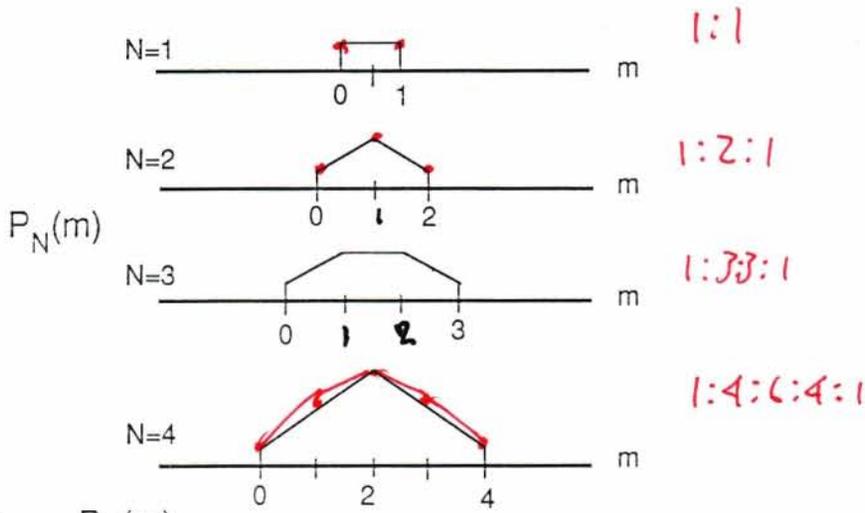


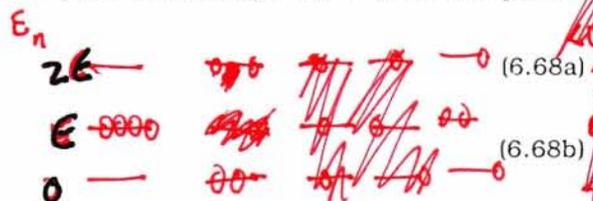
Figure 6.13 "Peaking" of the relative probability of obtaining a given fraction of tosses as heads, after N tosses of a single coin. As the number of tosses increases, the jagged shape (top diagram) becomes a smooth curve (bottom diagram), and one is more and more likely to obtain a result very close to the most probable result (namely, 0.5).

Example: Sorting particles into energy levels: Boltzmann distribution

In statistical mechanics, the fundamental problem is how to sort a fixed (large) number of particles, N_0 , into a known distribution of energy levels, with N_i particles in energy level, E_i , for a given total energy, E_0 , of all of the particles taken together:

$$\sum_i N_i = N_0$$

$$\sum_i N_i E_i = E_{TOTAL}$$



$$E_{TOTAL} = \sum_i N_i E_i = 2E \sum_i N_i = 2E \cdot 4$$

Handwritten notes and diagrams on the right side of the page. Includes a diagram of a particle in a box with energy levels and notes: '# LEVELS = 2 (1/2 NMR, E MR) ∞ (W/ ON SPRING, H-ATOM)'.

Let ~~in which~~ N_i and N_j represent the numbers of particles in energy levels ϵ_i and ϵ_j , and β is a constant. By comparing the classical and quantum-mechanical properties of an ideal gas confined in a box (see Problems), it is found that the constant, β , may be identified with $(1/kT)$: this statistical derivation is in fact sometimes used as a *definition* of temperature, as we will shortly find in the laser experiment. Equation 19-3 then becomes the "Boltzman distribution":

$$N_i/N_j = \exp[-(\epsilon_i - \epsilon_j)/kT]$$

$$N_i/N = \exp[-\epsilon_i/kT] / \sum_i \exp[-\epsilon_i/kT]$$

$$= (1/q) \exp[-\epsilon_i/kT]$$

$k = \text{Boltzmann constant}$
 $T = \text{Absolute temperature}$

$= e^{-\Delta\epsilon/kT}$ (19-4a)

$\Delta\epsilon = \epsilon_i - \epsilon_j$ (19-4b)

$q = \text{"partition function"}$ (19-4c)

Based on the Boltzmann distribution, a temperature is considered "low" when most of the particles have an energy much greater than thermal energy, $\epsilon_i \gg kT$, and a "high" temperature corresponds to particles whose energy is much less than thermal energy, $\epsilon_i \ll kT$. At "low" temperature, most of the particles will be in the lowest energy level(s), while at "high" temperature most of the energy levels will be equally populated, as shown schematically in Fig. 19-3. Of the energy-level situations we have examined so far, the "high" temperature condition applies to magnetic resonance experiments, while the "low" temperature condition is valid for infrared and optical-ultraviolet experiments, as shown in Table 19-1. In any case where

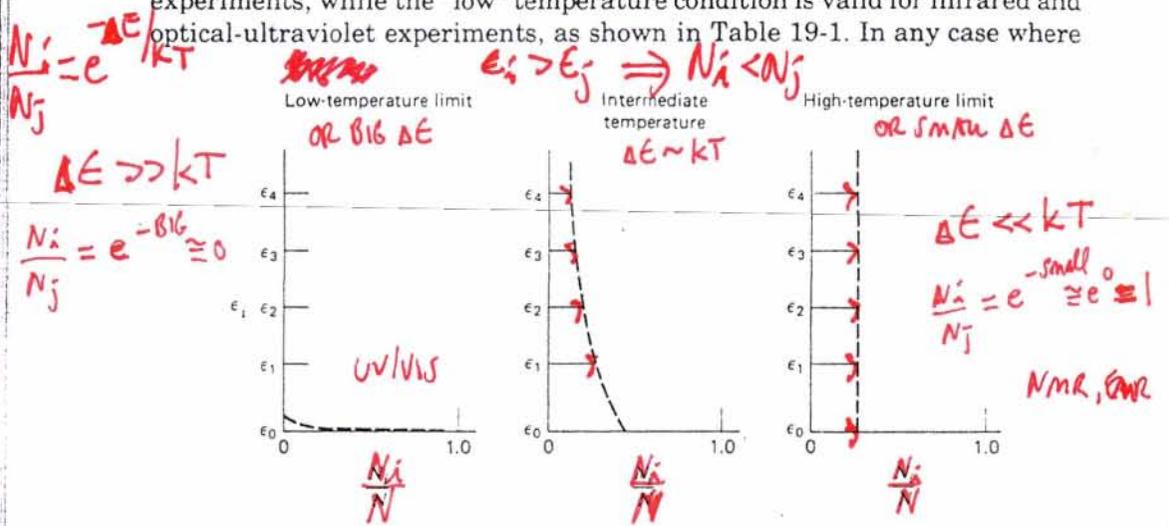
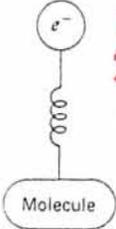
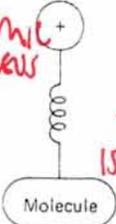
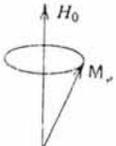
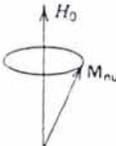
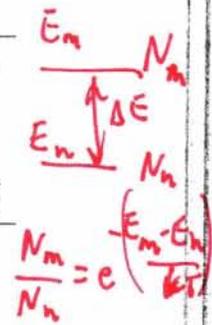


FIGURE 19-3. Relative equilibrium populations (abscissa) of various energy levels (ordinate) for three relative ratios of ϵ_i to thermal energy, kT . Low-temperature: $\epsilon_i \gg kT$; intermediate temperature, $\epsilon_i = kT$; high temperature, $\epsilon_i \ll kT$. Examples of each situation may be found in Table 19-1. Relative populations are obtained from Eq. 19-4a.

Table 19-1 Relative Populations of Two Adjacent Energy Levels, for Various Types of Natural Oscillations (See Eq. 19-4a)

Nature of Oscillation	Typical Natural Frequency of that Oscillation	Energy Difference ($\epsilon_1 - \epsilon_0$) Associated with that Frequency where $h\nu = (\epsilon_1 - \epsilon_0)$	Relative Populations of Lowermost and Next Highest Energy Levels, N_1/N_0 , for Various Temperatures		
			Room Temp. (295°K)	Liq. N ₂ (77°K)	Liq. He (4.2°K)
 <p>UV 200nm $\sim 1.5 \times 10^{15}$ Hz 10^{-18} J 2000Å = 1.5×10^{15} Hz 10^{-18} erg</p>			$10^{-106} \approx 0$	$10^{-404} \approx 0$	$10^{-7440} \approx 0$
 <p>I.R. C-Cl-C 1500 cm⁻¹ = 4.5×10^{13} Hz 3×10^{-20} J 3×10^{-19} erg</p>			0.00066	$7 \times 10^{-12} \approx 0$	$10^{-233} \approx 0$
 <p>65.8 GHz (electron spin resonance in magnetic field of 2.0 x 10^4 Gauss) 2.31T 4.4×10^{-23} J 4.4×10^{-24} erg</p>			0.989	0.960	0.471
 <p>100 MHz (proton) magnetic resonance in magnetic field of 23.5 Gauss 9.4T 6.6×10^{-26} J 6.6×10^{-26} erg</p>			0.99998	0.99994	0.9989



the system is at thermal equilibrium with its surroundings (i.e., almost all the time, for our purposes), there will be more particles in lower energy levels than in higher energy levels, as seen in Fig. 19-3.

As discussed further in the second half of this chapter, Table 19-1 partly explains why magnetic resonance absorption is so much weaker than infrared and optical absorption. Since the absorption intensity depends on the difference in the number of particles in the lower and upper energy