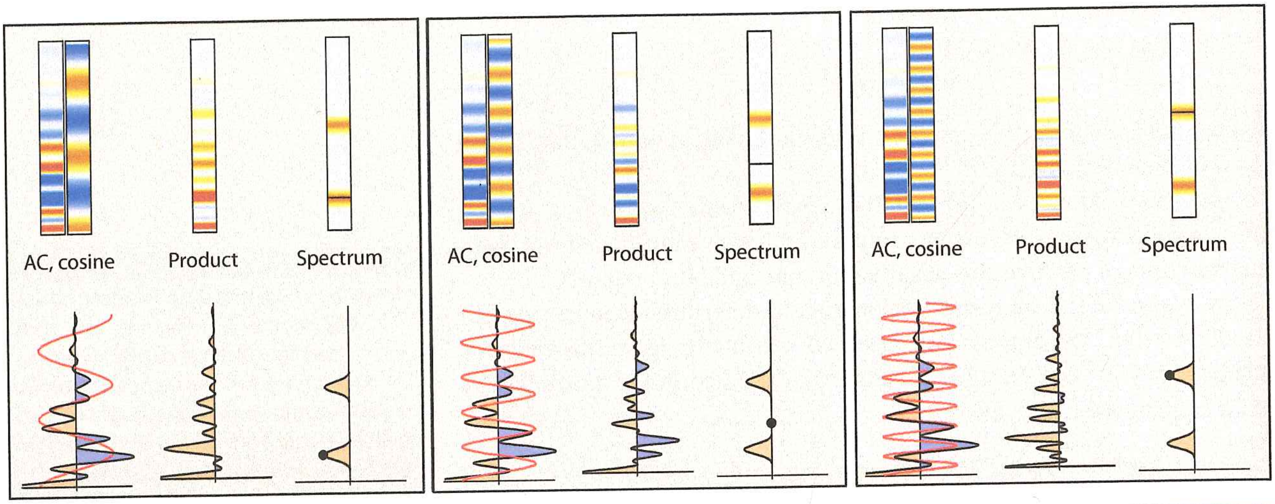

Figure 4.9
Fourier analysis of an autocorrelation function to yield the power spectrum. Three choices of frequency are shown; the black line or black dot reveals the frequency on the plots and also the value of the transform at that frequency. This spectrum consists of two peaks, both somewhat broadened. (As we will see shortly, the broadening corresponds to the decay of the autocorrelation in time, which is clearly visible.)

We have now seen the entire process of getting from signal to autocorrelation and finally to a spectrum. In figure 4.10, we show the process using color strip calculations from start to finish for a typical signal. This signal is a short burst, and the spectrum has turned out to be fairly broad and continuous. No finite set of cosines could be used to represent the autocorrelation, nor are there just a finite set of frequencies with power. This is why we needed the color strips and the Fourier analysis—to become facile with general and typical signals.

The Wiener-Khinchin Theorem

With the preceding discussion as pretext, the so-called Wiener-Khinchin theorem will come as little surprise: it states that the power spectrum and the autocorrelation are obtainable from one another; specifically, *the power spectrum is the Fourier cosine transform of the autocorrelation function.*

4.5 The Uncertainty Principle

Strictly periodic signals last forever and are perfectly correlated with themselves: they repeat at multiples of the period. At the other extreme, a signal might have no correlation with itself from moment to moment: knowing the signal at one instant says nothing about it the next; such a signal is said to correspond to *white noise* (noise that has power in all

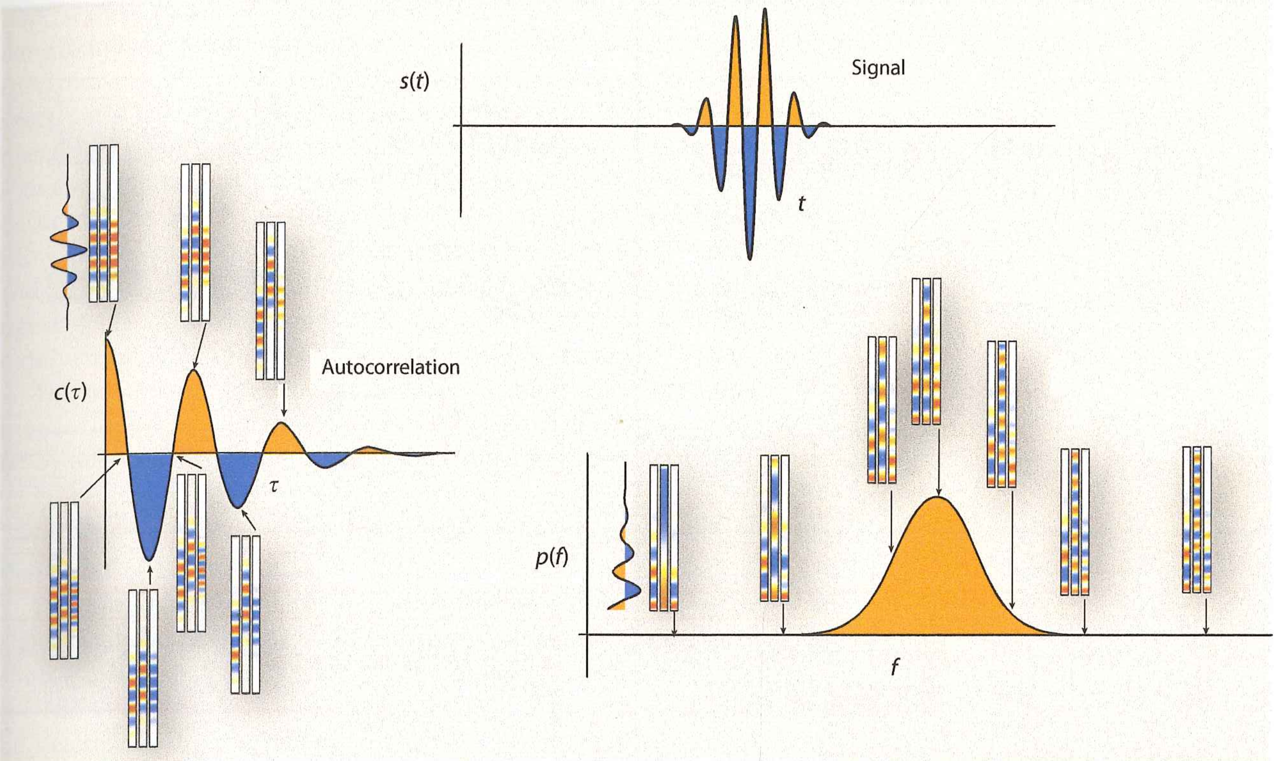

Figure 4.10
Construction of a power spectrum (bottom right), starting from the signal $s(t)$ (top), and passing through the autocorrelation $c(\tau)$ (bottom left). This signal is a short burst, but, apart from this, it is otherwise oscillating at a fixed frequency f_0. Starting with the signal, the autocorrelation is determined as in figure 4.7. From this, the power spectrum $p(f)$ is determined by multiplying the autocorrelation with cosines of different frequency and summing the result. In the lower right, the cosine is depicted in the middle strip of the groups of three. As the frequency increases to the right, the cosine oscillates ever more rapidly. The first two columns (the autocorrelation and the cosine) are multiplied together as before, the product is stored in the third column, and the average of the third column is taken to give the spectrum at the frequency of the cosine. Notice the evidence of the time-frequency uncertainty principle: the time data is a short pulse, and therefore the frequency spectrum is broadened.

regions of the audio spectrum). The autocorrelation for white noise has a peak only at $\tau = 0$ and falls immediately to 0 for $\tau \neq 0$.

Most sounds fall somewhere between white noise and perfect periodicity, often showing interesting structure in their autocorrelation. Their autocorrelation typically decays at longer times, indicating that the sound pressure at a given time is uncorrelated with the sound pressure far enough in the future.

The Fourier construction of the power spectrum reveals why a wide time pulse gives a narrow power spectrum, and vice versa. Choosing a cosine with a frequency shifted from the maximum of the power spectrum peak by $\Delta f = f_{max} - f = d$, we find (figure 4.11) that the product of

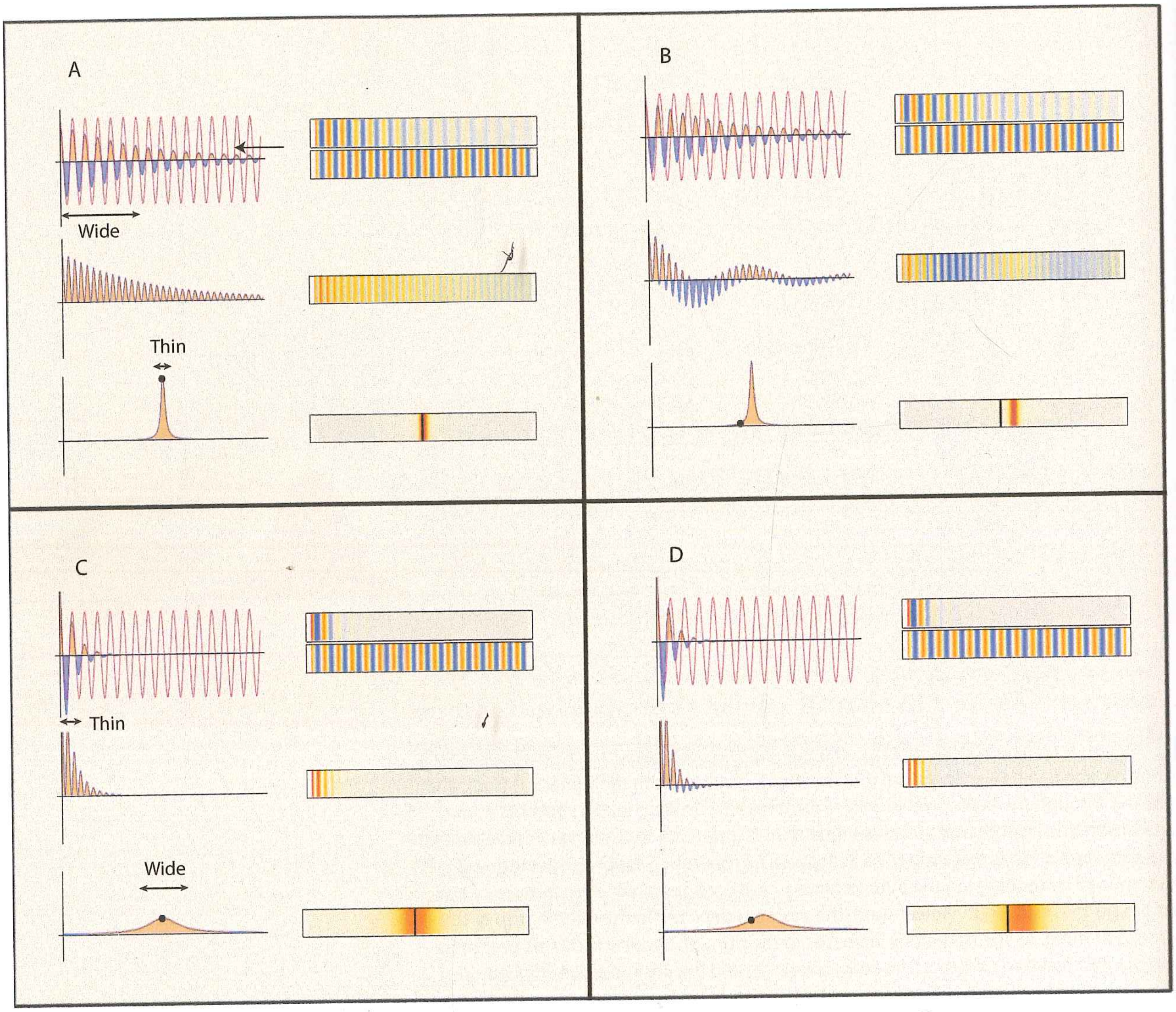

Figure 4.11
Two different autocorrelation functions are cosine-transformed into power spectra. A and B are an analysis of a (relatively) long-lived decaying autocorrelation at on-peak and off-peak frequencies. C and D analyze at the same frequencies but for a faster-decaying autocorrelation. Slow decay (broad in time) corresponds to a narrow frequency spectrum (panels A and B), whereas a fast decay (narrow in time) is paired with a broader power spectrum in frequency, in accordance with the uncertainty principle. Note that the off-resonant frequencies shown (cases B and D) correspond to reduced power, but the faster-decaying signal is reduced by less.

the cosine and the autocorrelation oscillates between positive and negative values with period $1/d$. This is quite apparent in panel B in figure 4.11. If the autocorrelation is decaying with time, the question arises, how many such beats are seen before the decay is significant? If the autocorrelation lasts several such beats (wide time signal), the negative and positive contributions will nearly cancel when forming the average, meaning the spectrum must be small at that frequency difference (narrow frequency spectrum). If at the same frequency difference d, the autocorrelation dies *before* the positive-negative cancellation can take place (narrow time signal), say, after just the initial positive part, the averaging out will not occur and the power spectrum will be larger (wider frequency spectrum). More power survives at off-center frequencies—that is, the spectrum is wider for a narrow or thin correlation function (see the annotations in figure 4.11). It is clear from this that *the shorter the duration of the autocorrelation, the broader the resulting power spectrum.*

The uncertainly principle states that the shorter the *time* signal, the broader the power spectrum in *frequency*. If we call the *time uncertainty* Δt, set by the autocorrelation duration, the frequency uncertainty Δf obeys

$$\Delta f \Delta t \sim 1, \tag{4.9}$$

that is, $\Delta f \sim 1/\Delta t$.

The uncertainty principle is one of the most important concepts in signal analysis. It thrusts itself into many situations, insisting that the shorter the time interval over which the sound is known or analyzed, the less is known about its frequency content. The time interval might be determined by the brevity of the recording or the window of time over which a recording is analyzed, but it equally well could be limited by the sound itself. If a trumpet player slides through a note in 1/20th of a second, it doesn't matter whether the sound file itself lasts much longer, because that note's frequency is actually not perfectly well-defined. According to the uncertainty principle, there is an intrinsic 20 Hz indeterminacy.

4.6 Autocorrelation and the Chorus Effect

No one mistakes a chorus for a soloist, and vice versa. Nor are 20 violins ever mistaken for a single violin. Yet if all singers are somehow in exact lockstep unison, with no vibrato or frequency drift, the sound some distance away would be indistinguishable from that attainable by a single voice.

Singers and violinists are never in perfect unison. Each has a slightly different pitch, which is drifting or oscillating (vibrato) slightly. The waveform cannot be perfectly periodic, even though it averages, say, 220 Hz in frequency. No law of physics dictates that we should hear a fused 220 Hz tone with a definite pitch in this circumstance, but we do. We hear a pleasant scintillation, but not individual singers, mistuned or not. This is the *chorus effect.*

If a soloist holds a 220 Hz note with no vibrato, the sound and the autocorrelation is perfectly periodic. Not so for a chorus: a somewhat random addition of sinusoids takes place, one for each partial of each singer's voice as crests and troughs from different singers drift in and out of phase since they are not perfectly in tune. A given peak in the signal, the sum of all singers, may not be at all predictive of where the peaks in the signal lie a quarter-second later, implying that the autocorrelation decays within that quarter-second (or perhaps much faster). The autocorrelation decay on a timescale Δt broadens the power spectrum by Δf, such that $\Delta f \, \Delta t \sim 1$, in accord with the uncertainty principle.

Software and hardware tools exist that mimic the chorus effect by creating and combining many slightly pitch-shifted instruments from an original single instrument sound. This can be done by hand, starting with a single track of a solo voice or instrument, randomizing its pitch by a percent or two (without changing the speed of playback—a nice trick made possible in many sound processing packages; see section 17.10) and combining it (mixing it) with the original. A convincing chorus effect emerges once this is repeated 5 or 10 times with slightly different pitches applied each time.

The decay of the autocorrelation is evident in just such a computer-generated tone, constructed from the addition of 40 slightly mistuned replicas of two voices separated by a perfect fifth interval (a ratio of 2:3 in frequency), each voice with four harmonic partials. The autocorrelation for a pair of unwavering perfectly tuned voices is seen in the righthand part of figure 4.12.

We usually take our cues for pitch from the first few prominent peaks in the autocorrelation, as we shall discuss at length in chapter 23. The first part of both autocorrelations, 80 singers and 2, is very similar—almost identical, in fact. Thus if autocorrelation peaks are indeed the key to pitch perception, the chorus and the duet must be judged to have the same pitch.

We can use the Fourier cosine calculation to understand these consequences, which shows that the power spectrum broadens when an autocorrelation decays rapidly; this is illustrated earlier in figure 4.11.

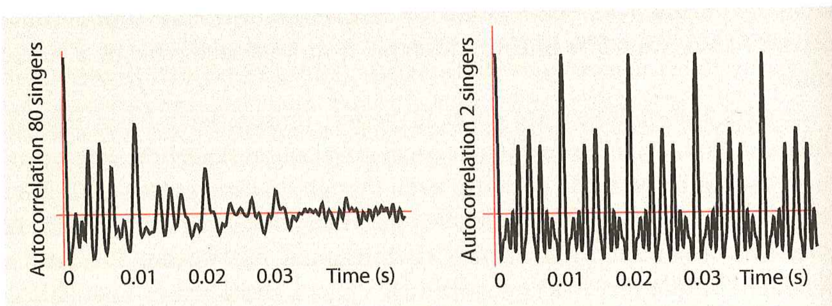

Figure 4.12

(Left) Autocorrelations of 40 slightly mistuned singers holding a 200 Hz note and another 40 singing a fifth above at 300 Hz, each slightly and randomly mistuned from the mean frequencies 200 and 300. (Right) Autocorrelations of one singer holding a 200 Hz note and another a fifth above, with no vibrato and perfect intonation. The chorus effect is associated with a decay of the autocorrelation function, on a timescale ΔT reflecting the range Δf of frequencies present, $\Delta f \Delta T \sim 1$.

4.7 Noise and Autocorrelation

We live in a noisy world. Sometimes noise dominates the sound. Can we extract the part that isn't noise? Dealing with noise is another great strength of using autocorrelation (and its partner, the power spectrum, which contains the same information). We examine two important cases that have implications for how our auditory system extracts information from noisy signals.

Autocorrelation and Fast Echoes

Suppose a sound is repeated with a short delay, of 0.05 to 0.001 s, corresponding to audio frequencies; such repetitions are echoes that are too fast to be heard separately. They do however "color" the sound, even imparting a pitch in some circumstances, as we will discuss further in sections 21.2 and 23.17. This applies to random white noise as well, repeated with a delay as follows: white noise is recorded or generated, and the signal is added to itself with a time delay, say, 5 ms, making a modified file. Clearly, if a large blip exists in the original file, the blip is now also found 5 ms later in the modified file. This gives a sharp peak in the autocorrelation function at 5 ms.

Masking Signals with Noise

Suppose we add some noise to a "clean" signal. The result is a ragged looking signal (figure 4.13). We may be able to hear the original clean signal buried in the noise even if we cannot visually detect its presence. The noisy

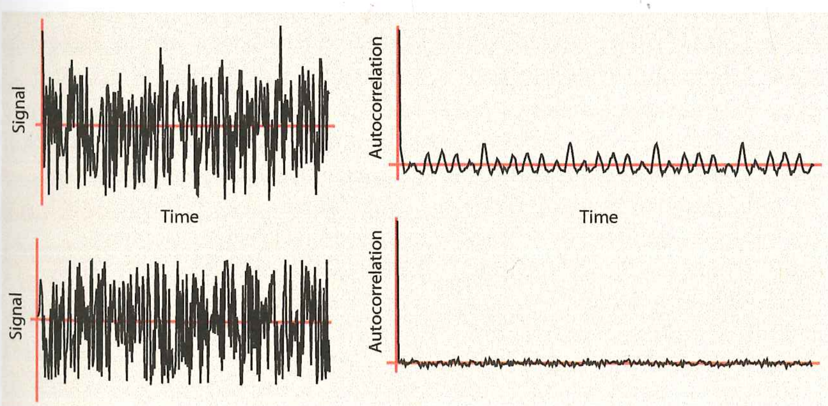

Figure 4.13
A pure noise signal and its autocorrelation are shown at the bottom. A noisy signal with a smooth periodic signal embedded (added to the noisy signal) appears at the top. The signal itself looks as ragged as the one at the bottom, which is pure noise. Computing the autocorrelation reveals the hidden periodic signal (top right).

signal s' is $s'(t) = s(t) + n(t)$, where $n(t)$ is the noise. The autocorrelation becomes

$$\begin{aligned} c'(\tau) &= \langle s'(t)s'(t+\tau)\rangle \\ &= \langle s(t)s(t+\tau)\rangle + \langle s(t)n(t+\tau)\rangle \\ &\quad + \langle n(t)s(t+\tau)\rangle + \langle n(t)n(t+\tau)\rangle \\ &= \langle s(t)s(t+\tau)\rangle + \langle n(t)n(t+\tau)\rangle. \end{aligned} \tag{4.10}$$

The last equality holds because there is no correlation between the signal and the noise—we assume they are completely independent. Cross terms such as $\langle s(t)n(t+\tau)\rangle$ average to zero and we can eliminate them.

The noise is random and decorrelates with itself immediately, leaving only the original signal. The autocorrelation is smooth in spite of the noise. This assumes, however, that enough data exist to do long averages; otherwise, the cross terms $\langle n(t)\, s(t+\tau)\rangle$ and $\langle s(t)\, n(t+\tau)\rangle$ do not average to zero.

This works so well that tiny signals can be rescued from a lot of noise, as seen in figure 4.13. A pure noise signal and its autocorrelation are shown at the bottom. A noisy signal with a smooth periodic signal embedded appears at the top. The combined signal looks as ragged as the one at the bottom, which is pure noise. However, computing the autocorrelation reveals the hidden signal (top right).

Box 4.2

Famous Fourier Transform Pairs

Time and frequency are the Fourier transform pair that concern us in this book. There are many other examples of Fourier transform pairs: quantities that are related by a Fourier transform. Anything that behaves as a wave has an uncertainty principle connecting Fourier transform pairs.

A famous Fourier transform pair involves quantum mechanics. Nature left the framers of quantum theory no choice: matter, mass, you and I, everything is ultimately a wave. The position of a particle (which is really a wave) and its velocity (momentum) are Fourier transform pairs, and are therefore subject to an uncertainty principle. This has the shocking consequence that the more certain we are about the position of a particle, the less we know about its velocity! The reverse is true too: the more we know about velocity, the less we know about position. This is known as the Heisenberg uncertainty principle. It reads $\Delta x \Delta v \geq \hbar/2m$, where m is the mass of the particle, Δx is the uncertainty in its position, Δv is the uncertainty in its velocity, and $\hbar$ is known as Planck's constant, which is a known number. Another example of a quantum uncertainty principle involves time and energy. In quantum mechanics, it takes a lot of time to pin down the energy of the system precisely. If we've followed a system for only a short time, then we are not very certain about its energy. That uncertainty principle reads $\Delta E \Delta t \geq \hbar/2$. Thus time and energy are Fourier transform pairs, as are position and velocity. The Einstein quantum relation $E = \hbar(2\pi f)$ relates the energy with frequency, most often the frequency of light emitted or absorbed in a quantum transition that changes the energy. Thus the time-energy uncertainty principle is really another time-frequency uncertainty principle in disguise.