1

Linear Systems Analysis

Linear systems analysis has found wide application in physics and engineering for a century and a half. It has also served as a principal basis for understanding hearing ever since its application to audition by Ohm (1843) and Helmholtz (1877). The successful application of these procedures to the study of visual processes has come only in the last two decades, but it has had a considerable impact on the field. It provides in fact much of the *raison d'être* of this book.

It is not our intent in this chapter to discuss the mathematical basis of linear systems analysis, but rather to give the uninitiated reader enough of a background and intuitive feel for the procedures to be able to follow those applications to visual problems which we will be discussing in this book. The reader who wishes to pursue these issues further is strongly urged to read any of a number of more comprehensive treatises on linear analysis (e.g., Bracewell, 1965; Goodman, 1968). An excellent short mathematical treatment of two-dimensional linear analysis and its application to visual problems is given in D.H. Kelly (1979a).

FOURIER THEOREM

Linear systems analysis originated in a striking mathematical discovery by a French physicist, Baron Jean Fourier, in 1822. He demonstrated that a periodic waveform of any complexity could be broken down (analyzed) into the linear sum of harmonically related sine and cosine waves of specified frequencies, amplitudes, and phases. (Harmonics are waves whose frequencies are integer multiples of some lowest, or fundamental, frequency). See Figure 1.1 for examples of sine waves. This is now known as Fourier analysis. Conversely, by the inverse Fourier transform, any desired periodic waveform could be built up (synthesized) by adding together harmonically related sine waves of appropriate frequency, amplitude, and phase. Later, it was shown that any nonperiodic waveform could also be described as the sum of sine waves if all frequencies (not just harmonically related ones) were included.

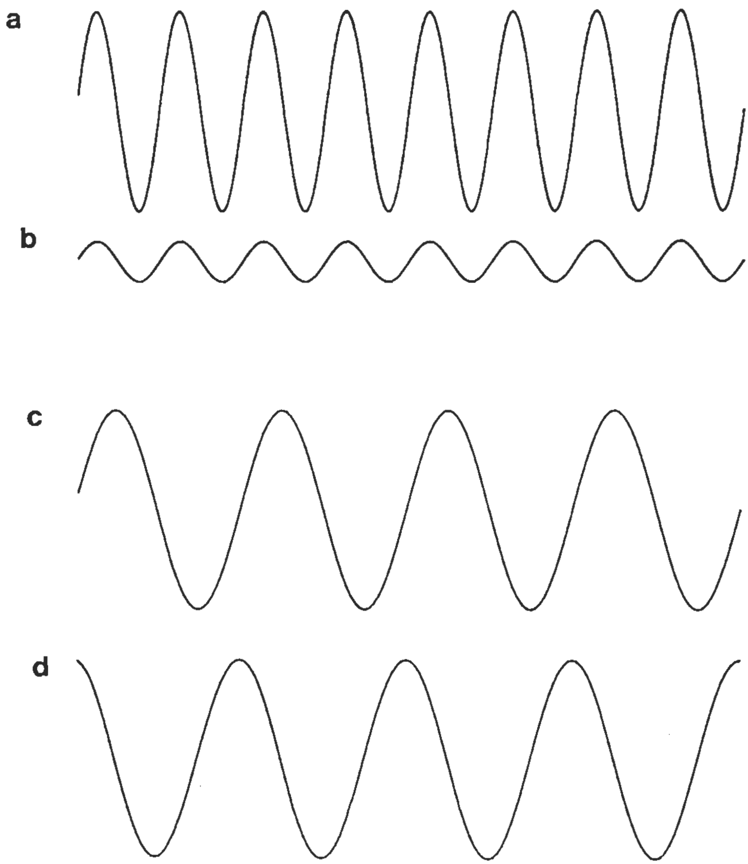

FIG. 1.1 Examples of sine (and cosine) waves. The waveforms in A and B have a spatial frequency which is twice that of the waveforms in C and D. The sine wave in A is equal in frequency and phase to that in B, but its amplitude is 5 times as great. The waveforms in C and D are equal in frequency and amplitude but differ by 90° in phase.

Formally, the Fourier transform of a spatial (or temporal) waveform is

$$F(\alpha) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi\alpha x}dx$$

where x is the location in space (or time, if it is a temporal waveform), α is spatial (or temporal) frequency, and $j = \sqrt{-1}$. The inverse Fourier transform from spatial or temporal frequency to space or time is

$$F(x) = \int_{-\infty}^{\infty} f(\alpha)e^{j2\pi\alpha x}d\alpha$$

The usefulness of Fourier analysis is largely that it provides a way in which complex waveforms or shapes can be described and dealt with quantitatively. One can, of course, specify a complex waveform such as that shown in the top of Figure 1.2 by its amplitude at each point in space: it is "so high" at this point, and "so high" at that point, etc. But there is little that one can say in summary about it except that it is wavy and has several bumps. It would be difficult to make any but the vaguest comparisons of this to some other similar waveform. Breaking it down into elemental units with Fourier analysis, however, allows

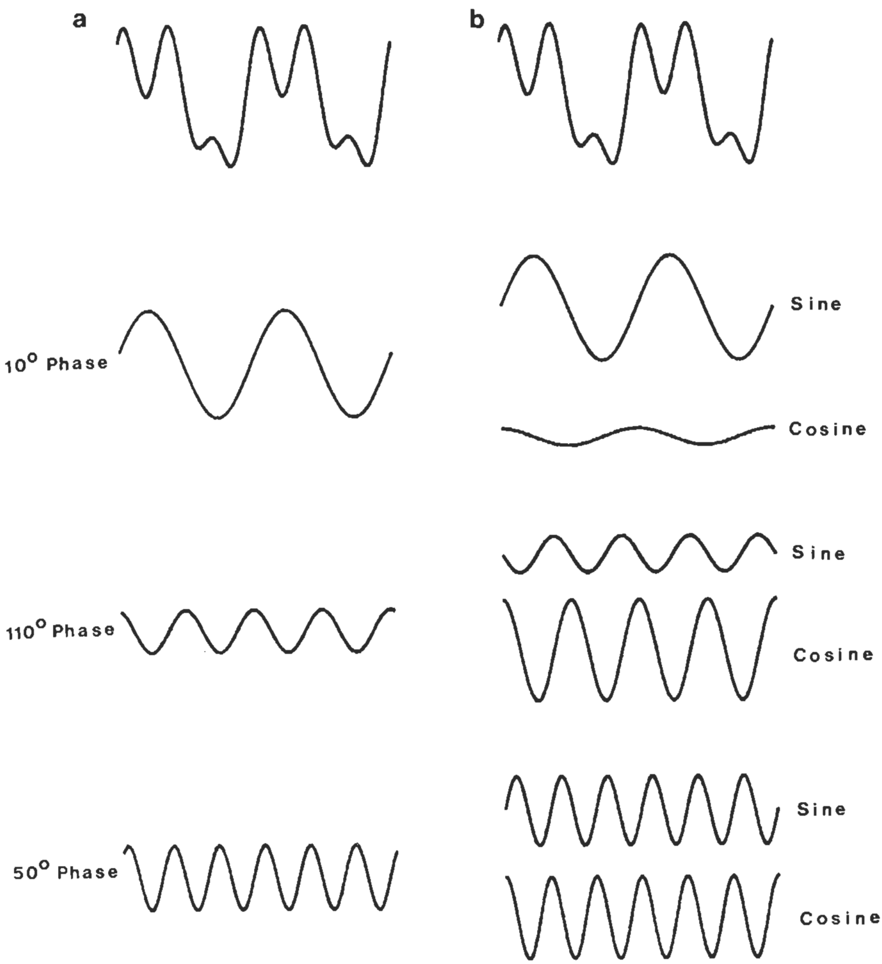

FIG. 1.2 The arbitrary waveform shown at the top can be analyzed into (or synthesized from) either the sine waves of specified frequency, amplitude, and phase shown in A or, alternatively, the three pairs of sine and cosine waves of specified frequency and amplitude shown in B.

one to specify it quantitatively and compare it to other waveforms in an economical way.

A second aspect of Fourier analysis that makes it very useful in many applications is that any waveform can be specified by combinations of a limited number of elemental components, the number depending on the degree of precision required. Sine waves thus provide a universal representation for complex waveforms.

Sine Waves

Examples of sine waves, the basic elements in Fourier analysis, are shown in Figure 1.1. We are most accustomed to dealing with such sinusoidal oscillations in the time domain, e.g., a rapid oscillation of air pressure, a sound wave, produced by a tuning fork or a bell. In our consideration of spatial vision, however, we shall be more concerned with oscillations across space, e.g., ripples on a lake.

The frequency of a sine wave is the number of oscillations per unit distance or time. The patterns in Figures 1.1A and B are of twice the frequency as those in Figures 1.1C and D. So a certain *temporal frequency* might be 3 cycles (or complete oscillations) per second, commonly abbreviated Hz (after the physicist Hertz, and so pronounced). A particular *spatial frequency* (an oscillation of luminance or color in space) might be 3 cycles per centimeter. Since the dimensions of visual stimuli are much more usefully specified in terms of the angle subtended at the eye (visual angle), the usual specification of spatial frequency in vision is in cycles per degree visual angle (c/deg), as described in Chapter 2.

The amplitude of a sine wave is the distance from peak to trough of the wave divided by 2. The sine wave in Figure 1.1B is the same frequency and phase as that in Figure 1.1A, but is of ⅓ the amplitude. For a more complex waveform the amplitude of the overall pattern is not necessarily the same as the amplitude of the fundamental or any other harmonic component. For instance, the amplitude of the sine wave fundamental in a square wave pattern (see Figure 1.3) is greater than that of the pattern as a whole.

The contrast of a visual pattern is related to the amplitude in the sense that both are measures of the height of the waveform, but it is important to distinguish between them. The *contrast* of a visual pattern is conventionally given by the Michelson contrast:

$$\frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}$$

where $L_{\max}$ and $L_{\min}$ are the luminances of the peak and trough of the waveform, respectively. As can be seen, then, contrast is a relative measure. Waves are sometimes described not by their amplitude but by their power, which is amplitude squared. The power spectrum is the power at each of the various frequencies, irrespective of phase.

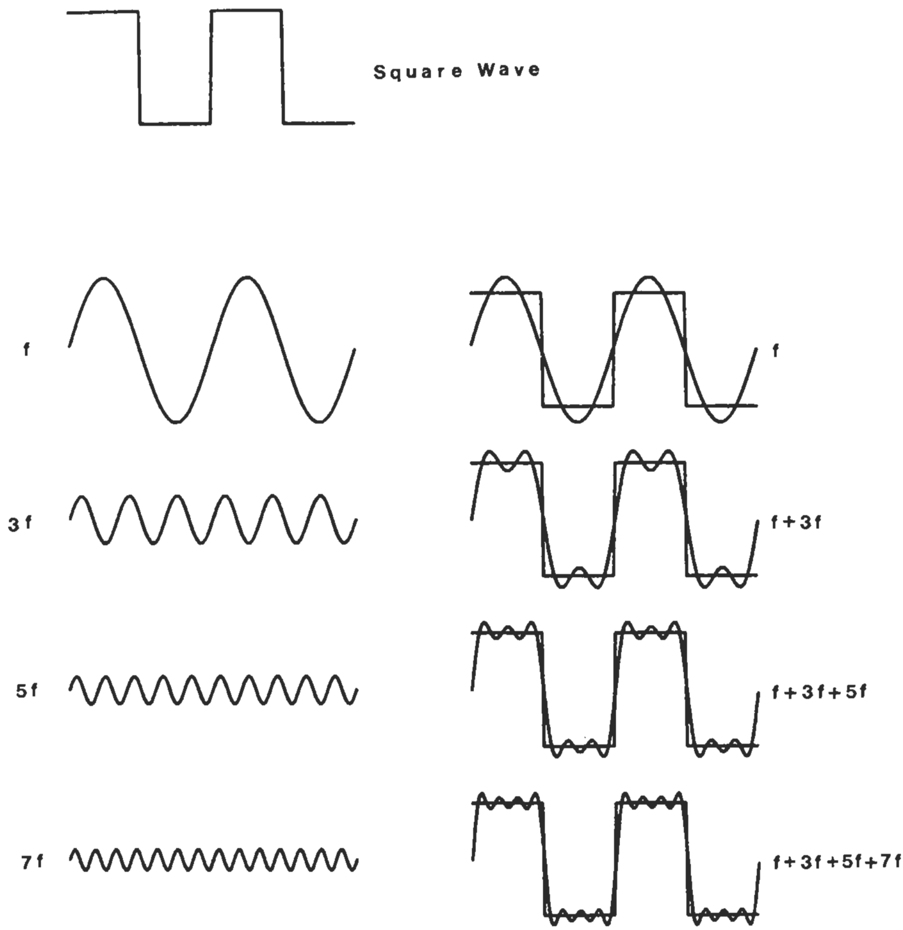

FIG. 1.3 Analysis of a square wave into its components. On the left are shown a section of a square wave and its first four components in appropriate relative phases and amplitudes. On the right is shown the superposition of these components on the original square wave.

The third critical variable is the phase, which refers to the position of the sinusoidal wave with respect to some reference point. In Figure 1.1D, the sine wave is of the same amplitude and frequency as that in Figure 1.1C, but it is shifted 90° in phase. One can distinguish between two different usages of the term "phase" in describing sinusoids. We will refer to these as absolute and relative phase, although strictly speaking all phase must be relative. By "absolute phase" we will mean the position of a sinusoidal wave with respect to some particular fixed spatial location. By "relative phase" we will refer to the relative phase angles (differences in absolute phase) among the multiple frequencies in a pattern. A sine wave and a cosine wave are identical except that one is shifted 90° in phase with respect to the other. The sine wave has a zero at the spatial location used as a phase origin, while the cosine wave has a maximum at that point. Since a sinusoidal wave of a given phase can be expressed as an appropriate combination of sine and cosine components of fixed (0) phase, an analysis of a complex waveform in terms of sines and cosines is equivalent to one utilizing sine waves of varying phase. In many ways it appears that using sine and cosine components is more similar to the way in which the visual system itself might deal with phase, a possibility we shall discuss later.

Examples of Fourier Analysis

An illustration of Fourier synthesis and analysis is shown in Figure 1.2A. The arbitrary waveform shown at top can be "synthesized" or built up from the particular three sine waves, specified in frequency, amplitude, and phase, shown below at the left. That is, if at each point the values of the three sinusoids are added together linearly, their sum will produce the top waveform. In the reverse process, one can decompose (analyze) the top waveform into the sum of the three sine waves below. In Figure 1.2B is shown the analysis of the same waveform into sine and cosine components (of fixed phase), with the amplitudes being specified for each of the three frequencies. Again, the sum of these (six) sinusoids adds up the complex waveform at top.

In Figure 1.3 we show another example of a complex waveform analyzed into its component sine waves. This waveform is called a square wave, and it is of particular importance to vision for several reasons. It is a not uncommon distribution of light, being, for instance, what one would obtain from running a photocell across a picket fence. A square wave also has sharp edges, which have often been thought to be of particular importance for vision. Further, square wave gratings have been widely studied in psychophysical and physiological experiments.

By the use of Fourier analysis, a square wave, which is a periodic function, can be broken down into harmonically related sine waves. The fundamental sine wave component, f, is of the same frequency and phase as the square wave itself, but it has an amplitude which is $4/\pi$ times that of the square wave. If it is added to a sine wave of the same phase but of 3 times the frequency ($3f$) and ⅓ the amplitude as the fundamental, the resulting waveform can be seen to be closer to the square wave. The addition of $3f$ tends to sharpen up the edges a bit, and to cut back the too large peaks and troughs of the original sine wave. As one continues to add odd harmonics in decreasing amplitude at appropriate phases, the resultant waveform more and more closely approximates that of a square wave. So a square wave of frequency f and amplitude 1 can be analyzed into the sum of sine waves which are the odd harmonics of f (i.e., odd integral multiples). Specifically, it is equal to

$$4/\pi\ [\sin(f) + \sin(3f)/3 + \sin(5f)/5 + \sin(7f)/7 + \ldots + \sin(nf)/n].$$

One way to grasp the nature of Fourier analysis is to consider it as a correlational process. In the Fourier analysis of a waveform, one is specifying the extent to which this waveform is correlated with (or overlaps) sine waves of different frequencies and phases. A sine wave is obviously perfectly correlated with a sine wave of the same frequency, amplitude, and phase and can thus be specified by a single point in the frequency domain. A square wave is highly correlated with a sine wave of the same frequency, f, completely uncorrelated with a sine wave of frequency $2f$, partially correlated with a sine wave of $3f$ (but less so than with f), etc. Its Fourier spectrum, the extent to which it is correlated with sine waves of various frequencies, thus consists of f and its odd harmonics in decreasing amplitude.

Filters can be considered as devices that carry out such a correlation. A filter passes information insofar as the input matches, or correlates with, the filter function. Imagine a bank of filters tuned, respectively, to f, $2f$, $3f$, etc. A sine wave of frequency f would activate just the f filter; a square wave would activate the f filter and also the $3f$ and $5f$ and . . . filters, to decreasing extents.

The Fourier amplitude spectrum of a waveform can be graphically represented in various ways. Two useful ones are shown in Figure 1.4. In Figure 1.3, an example of a spatial square wave was shown. The representation of this in the frequency domain is shown in Figure 1.4A, with frequency along the x-axis and amplitude along the y-axis. Another form of frequency representation (which in general is more useful since it can also be used to show the spectrum of two-dimensional patterns) is a polar plot, as shown in Figure 1.4B. Here, spatial frequency is indicated by the distance out from the center and orientation by the polar direction. Amplitude, in such a plot, is often arbitrarily shown by the size or intensity of the spot at a particular frequency and orientation. (Neither plot shows the phase of the components.) Such a two-dimensional plot is similar to the actual two dimensional power spectrum of a pattern obtained by a simple optical diffraction technique (Goodman, 1968).

It can be seen that in terms of its Fourier spectrum, a square wave is a rather complex waveform. Paradoxically, a single narrow bar is an even more complex waveform: since it is a nonperiodic pattern, its synthesis requires not just harmonically related sine waves, but contributions from a continuous spectrum of frequencies across a wide range. In fact, it is by no means intuitively obvious that one could find any combination of sine waves that would add up to a single bar on a uniform background, but it can be done by adding together many sine waves that are in phase at this one location yet cancel everywhere else.

Frequency Domain Versus Space Domain

We pointed out above that a bar, strictly localized in the space domain, has a broad, continuous spectrum in the frequency domain. This raises the general question of the relationship between the frequency and space domains in Fourier analysis. Recall that one can go back and forth between the space domain

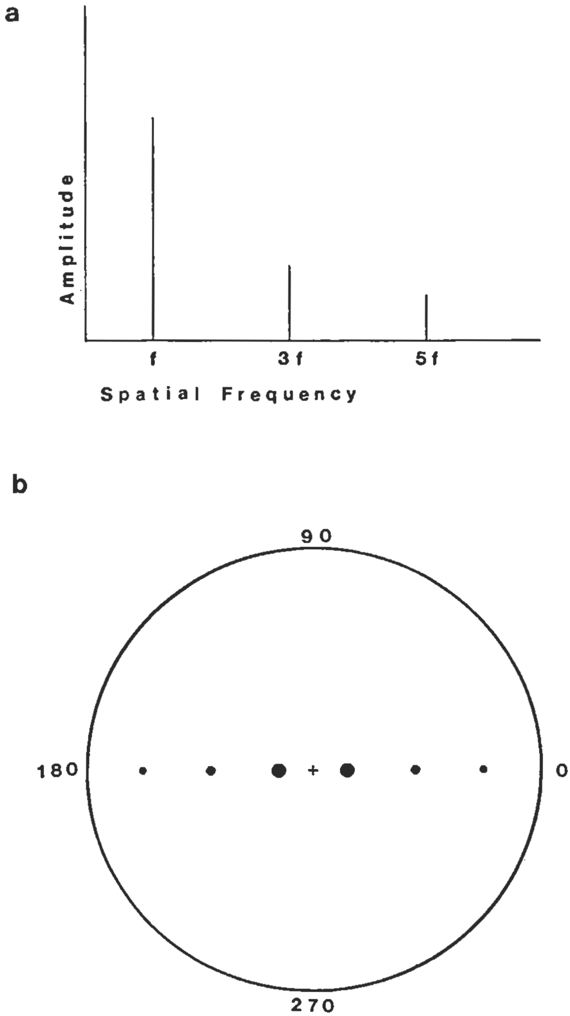

FIG. 1.4 Two graphical representations of the first three components of a square wave. In A, frequency is plotted along the x-axis, and amplitude along the y-axis. Phase is not represented. In B, the spatial frequency corresponds to the distance from the center, and orientation to the polar direction. Amplitude is here represented by the diameter of the spot. This would be the spectrum of a *vertical* square wave.

and the frequency domain representations by the Fourier transform and the inverse Fourier transform, respectively. There are some interesting, contrasting relationships between these two domains. The simplest, most discrete pattern in the frequency domain, a single spatial frequency, corresponds to a sine wave grating in the space domain that covers all of space, extending infinitely in all directions. On the other hand, the Fourier transform of a single, infinitely small, discrete spot in space covers the whole two-dimensional spatial frequency spectrum, since it contains power at all orientations as well as at all spatial frequencies. So a spot of light, chosen by many visual scientists in an attempt to use the simplest stimulus, is from the point of view of Fourier analysis one of the most complex patterns of all! The spatial frequency spectrum of a spot is very similar to the temporal frequency spectrum of a click. A spot of light and an auditory click appear in the space and time domains, respectively, to be elementary stimuli, but in the spatial and temporal frequency domains they are very extensive, complex patterns. A single spot has very little power at any spatial frequency and orientation. (It is interesting that a large array of randomly positioned dots, e.g., the night sky full of stars, has essentially the same power spectrum as a single dot, but with much more power at each spatial frequency, and a different phase spectrum.)

Sampling Theorem

It is a theorem of euclidean geometry that two points determine a line. This is readily grasped intuitively. A somewhat less obvious geometric finding is that three points determine a circle. The question we wish to raise here is how many points determine a complex waveform. How often must one sample the waveform in order to specify its frequency components? The answer is that with a certain number of appropriately spaced point samples, N, all frequencies below $N/2$ can be specified. So, for instance, 30 regularly spaced samples across 1° visual angle will specify the frequency, amplitude, and phase of all frequencies below 15 c/deg. Thus only slightly more than two samples per cycle of the highest frequency present are required to specify the contributions of all sine waves in a complex pattern. Further data points add no additional information. This is known as the sampling theorem.

The sampling theorem has relevance to the study of the visual system in considering the number and spacing of the visual receptors and other elements in relation to the resolving power of the system, as discussed further in Chapter 2. If the optics of the eye are such as to pass patterns of only, say, less than 60 c/deg, there would be no point in having more than 120 receptors per degree visual angle. The argument can also be put the other way: if receptors can only be packed in at 120 per degree visual angle, there is nothing to be gained from developing a better optical system than one which cuts off at 60 c/deg. This discussion only considers one-dimensional patterns, and the retina is of course a two-dimensional structure. The optimum sampling of such two-dimensional structures is a somewhat more complex question (Crick, Marr, & Poggio, 1981), but the results do not significantly change the conclusions stated above.

Nonsinusoidal Basis Functions

The Fourier theorem states that any waveform can be analyzed into its component sine waves. Sine waves, then, are the elementary units, or basis functions, in this analysis. It is also possible to carry out analogous analyses utilizing other basis functions. For instance, in the Hadamard transform, waveforms are analyzed into square waves rather than into sine waves. So Fourier analysis into sine waves is by no means the only possible linear decomposition procedure. However, sine waves have certain properties that which make them attractive as a basis function, particularly for analyzing optical and visual systems. Sinusoidal waves are of interest and importance because spatially invariant linear systems do not alter their shape. Lenses, and perhaps also certain arrays of visual cells, are to a first approximation such linear systems. For example, a sinusoidal pattern of light passed through a lens with aberrations is not changed in waveform or shape but merely in amplitude and possibly in phase. It is still a sine wave. Thus, even with imperfect optics, no additional frequencies are introduced for sine wave basis functions. That is not the case, however, for square wave basis functions. For these reasons, we will discuss here only Fourier analysis, and only sine waves as a basis function. (Systems with certain symmetries, e.g., two-dimensional circular symmetry, may also pass combinations of sine waves, for instance Bessel functions, without change of shape.)

APPLICATIONS TO VISION

The Fourier theorem can be applied to any sort of oscillation and to variations in either time or space. It has found wide application to many sorts of problems. For instance, in audition one deals with oscillations of sound pressure in time. A physicist might be concerned with variations in heat across space or in time (Fourier in fact developed the mathematics named after him as a tool to study heat waves). A meteorologist might apply Fourier analysis to the study of oscillations in rainfall or temperature across the year. In vision, we are concerned with variations in both time and space, and with variations in either the intensity or the wavelength of light.

Consider, for instance, an ordinary visual scene. If you were to run a photocell across the scene at some constant level, say in a horizontal line from left to right, the output of the photocell would go up and down as it encountered light and dark areas. The result would be a complex waveform related to the spatial pattern of light across the scene. Correspondingly, if such a pattern were imaged on the retina, the output of an array of receptors would constitute samples of this complex waveform. In its application to vision, then, the Fourier theorem and its extension states that these, like any other waveforms of variations in light across space or time, could be analyzed into the sum of sine wave components, or synthesized by adding together sine waves of appropriate frequency, amplitude, and phase.

In applying Fourier analysis to vision, we can first consider a one-dimensional pattern, such as the grating shown at top in Figure 1.5. If this pattern were of

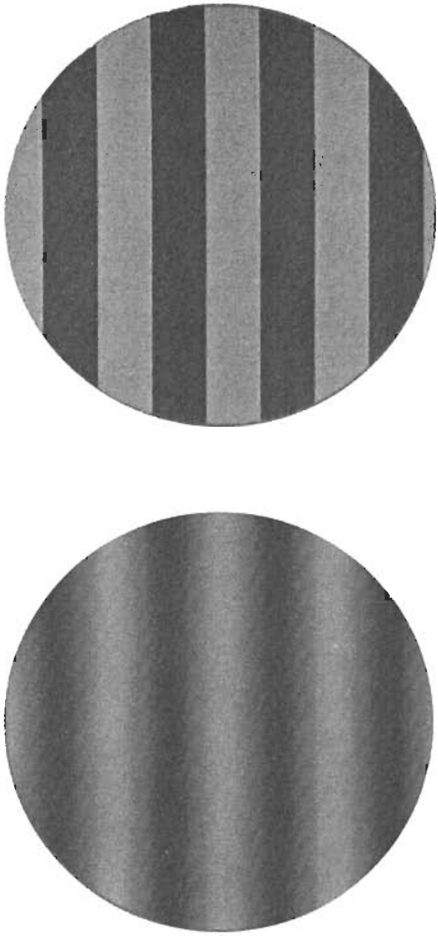

FIG. 1.5 Square (top) and sine wave (bottom) gratings of the same spatial frequency.

infinite height, there would be a variation in light only from left to right across the pattern, not from up to down. We can thus consider it a one-dimensional pattern, although of course any real object terminates in the other dimension as well and is thus really two-dimensional. The shape of the waveform that would result from running a photocell with a narrow acceptance angle across such a pattern and plotting the light intensity would be a square wave such as was shown in Figure 1.3. This pattern is thus called a square wave grating. As we discussed earlier, a square wave is really a fairly complex pattern, from the point of view of Fourier analysis, being made up of an infinite series of odd harmonics of a sine wave of the same frequency as the pattern. The simplest one-dimensional spatial stimulus, then, would be a grating pattern like that shown at the top of Figure 1.5, but with a sinusoidal variation in light across it from left to right. The luminance, L, at each point in x across this pattern would then be

$$L(x) = L_m\,[1 + c \sin(2\pi f x + \phi)]$$

where L_m is the mean luminance, c is the contrast, f is the spatial frequency, and ϕ is the spatial phase. Such a *sine wave grating* is shown at bottom in Figure 1.5. By the fundamental theorem of Fourier analysis, then, any one-dimensional visual pattern could be synthesized from the sum of such sine wave gratings of the appropriate frequencies, amplitudes, and phases.

In the various waveforms we have been discussing and whose Fourier spectra we have diagrammed, we have implicitly assumed that the values vary from positive to negative across the waveform, with a mean of zero. Such would have to be the case, for instance, for a pattern (a true sine wave) to have all its power at just one frequency. Since negative light does not exist, however, any visual pattern must consist of variations in light about some mean level. This introduces a zero spatial frequency, or DC component, into any actual visual pattern, the DC level being the mean luminance level of the pattern.

Two-Dimensional Fourier Analysis

The actual retinal image of most scenes varies not just from left to right, as in the grating patterns we have been showing, but in the vertical dimension too. Ignoring depth, the visual world is two-dimensional, as is the retinal image. Two-dimensional patterns can be dealt with by the same procedures of Fourier analysis, except that the orientation of the sine wave components must also be specified. Any two-dimensionally varying pattern can be analyzed into sinusoidal components of appropriate spatial frequency, amplitude, phase, and *orientation*.

Formally, the two-dimensional Fourier transform is

$$F(\alpha,\beta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-j2\pi(\alpha x + \beta y)}dx\,dy$$

and the inverse transform is

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(\alpha,\beta)e^{j2\pi(\alpha x + \beta y)}d\alpha\, d\beta$$

where x and y are spatial coordinates, and α and β are horizontal and vertical frequencies.

On a polar plot a vertical grating, such as that shown at top left in Figure 1.6, has all its power along a horizontal axis, since the pattern varies only horizontally, see top right in Figure 1.6. Correspondingly, a horizontal grating has its power along the vertical axis. The simple, two-dimensional plaid pattern shown at bottom left in Figure 1.6 was made by the addition of the vertical grating shown at top plus a horizontal square wave grating of the same frequency. It thus has energy at each of these orientations (see bottom right). Also shown, at middle left in Figure 1.6, is a checkerboard pattern. This also is made from a vertical and a horizontal square wave grating, but in this case it is the mathematical *product* of the two rather than their sum. The *Fourier fundamentals* are thus on the diagonals, at 45° with respect to the original gratings, and the spatial frequencies of the fundamentals are $\sqrt{2}$ times that of the grating fundamentals. The numerous higher harmonics, cross products of the various fundamentals and higher harmonics, are at a variety of other orientations and frequencies (middle right). It is thus apparent that there are vast differences between the usual description of a checkerboard in the space domain (vertical and horizontal edges; one particular bar or check width) and in the frequency domain (no power at vertical or horizontal; various spatial frequency components at a variety of orientations).

Two-dimensional spatial frequency analysis and synthesis is of course not limited to simple repetitive patterns such as checkerboards. In Figure 1.7 are photographs of the progressive synthesis of a very complex non-repetitive pattern. At top left is shown a grating of the frequency and orientation whose amplitude in the figure is the largest (a very low frequency vertical grating), with its power spectrum alongside. At top right is the sum of the two largest components in the pattern (the next largest component being a horizontal low frequency), with again their power spectra alongside, etc. By the time 64 components have been summed together, the shape of the pattern becomes recognizable, and 164 frequencies are enough to easily make out the specific picture portrayed. It is unlikely that the pattern could be so economically encoded in the pure space domain, with just 164 points (the original picture had some 65,500 points). Note also that the largest frequency components are all at low spatial frequencies. Just a small number of these low frequency components contain enough information to identify the picture.

Local and Global Analysis

Fourier analysis is in principle a completely global process. That is, the Fourier amplitude at a given spatial frequency and orientation is determined by the waveform at that orientation extending infinitely in space. In the case of a temporal Fourier analysis, one is dealing with infinitely extended time. However, in applying Fourier analysis to any real-life situation it is clear that such completely global analyses cannot operate, since measurements in neither time nor space can extend to infinity. So any practical application of Fourier analysis must

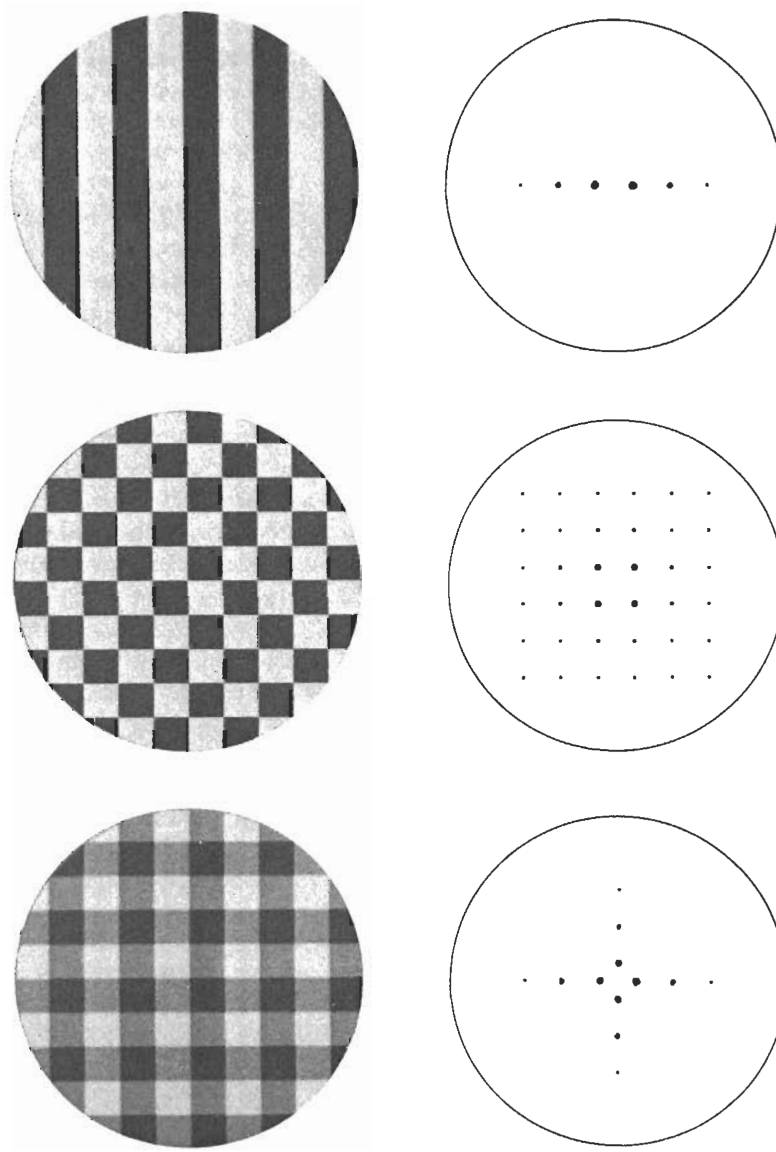

FIG. 1.6 Square wave grating, checkerboard, and plaid patterns. On the left are three related patterns: (top) a square wave grating, (middle) a checkerboard produced by the *cross multiplication* of vertical and horizontal square waves of this frequency, and (bottom) a plaid that results from the *addition* of vertical and horizontal square waves. On the right are polar plot representations of the spatial frequency spectra of these patterns.

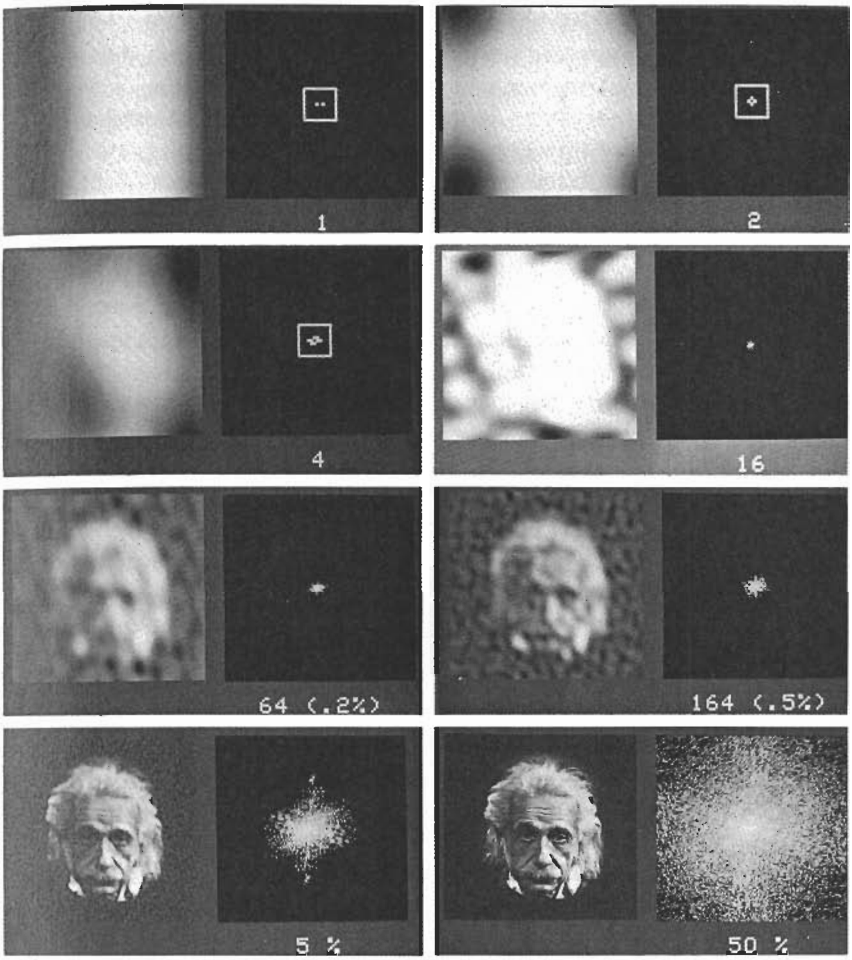

FIG. 1.7 Successive steps in a Fourier synthesis of a complex pattern by the addition of progressively greater numbers of those components which have the highest amplitudes. Alongside each pattern is its power spectrum, the central part of which has been enlarged for each of the first three combinations, for ease of seeing which frequencies are involved. See text for further explanation. We are grateful to Eugene Switkes for preparing this figure. (Derived from a photograph kindly provided by the American Federation of Teachers, with permission for its use).

involve an approximation, dealing with waveforms that must be at least somewhat localized in time and space.

In the applications of Fourier analysis to vision, the question of global versus local analysis refers to still further limitations on the process. Visual space is not infinite, but the visual field does extend some 200° or more; the duration of a single visual event, or visual time, is not infinite, but it could potentially extend for minutes or hours. It appears, however, that in analyzing spatial and temporal variations the visual system fragments patterns much more than is demanded by the spatial extent of the retina or the length of the day. Temporal events are analyzed into small time segments, and space into fairly small parts. Thus, in applying Fourier analysis to the operation of the visual system, or in considering the question of whether the visual system may itself utilize something like Fourier analysis in solving its problems, only a fairly localized analysis should be considered. What we shall be concerned with, then, is something intermediate between a totally global Fourier analysis and a totally localized point-by-point analysis. We can think of this intermediate condition as being a spatially and temporally patchwise approximation to a Fourier analysis.

A sine wave of an infinite number of cycles has a single frequency. However, if the number of cycles is curtailed somewhat (as it is in any practical application), it no longer has a spectrum consisting of one point, but rather of a group of nearby frequencies: its spectrum is somewhat broadened. The smaller the number of cycles, the broader will be the band of frequencies in its spatial spectrum. The breadth can be described by its bandwidth: the distance between the points at which the amplitude falls to half its maximum. This is usually specified in octaves (an octave being a 2/1 ratio).

A filter that correlated the input pattern with a sine wave of infinite extent would have absolutely precise spatial frequency information, since the only thing it would pass would be a sine wave of that precise frequency. If the filter contained only a limited number of cycles, it would pass a band of frequencies, and thus there would be some uncertainty in interpreting its output. A given output could have been produced by any of several nearby frequencies within its passband.

Although an extended sine wave filter would have very precise spatial frequency information, it would have considerable uncertainty with respect to the spatial location of a spatially delimited pattern. That is, a spatially restricted pattern that contained spatial frequencies within the filter's passband could be located at any of several positions within the field and produce the same output from the filter. If the filter were of limited extent (had a limited number of cycles), it would have a broader spatial frequency passband, but it would have reduced uncertainty about spatial location. There is thus a trade-off between uncertainty with respect to spatial frequency and uncertainty with respect to spatial location. A similar situation holds in the auditory system between specification of temporal frequency and the time at which an event takes place. An extended temporal frequency sample gives good temporal frequency resolution at the expense of information about time of occurrence, and vice versa. This problem was examined by Gabor (1946), who showed that a temporal filter with a gaussian fall-off (in both time and frequency, since a gaussian shape is a Fourier transform of itself) would minimize the product of the time and frequency uncertainties. A large number of cycles within the envelope would produce little frequency (but much time) uncertainty; a limited number of cycles would produce little time (but much frequency) uncertainty. But the total uncertainty would remain constant, given an overall gaussian envelope.

A further aspect of the Gabor functions (localized frequency filters with gaussian envelopes) is that, like pure sine and cosine waves, they can provide a complete description of any complex waveform (Helstrom, 1966). These functions can therefore give the concept of a spatially localized spatial frequency analysis, such as we were discussing earlier, mathematical precision. Although Gabor was considering one-dimensional temporal frequency analysis, his equations can equally well be applied in the spatial frequency domain (Marcelja, 1980), and expanded to two-dimensional patterns, with a gaussian taper in the second (orientation) dimension as well (Daugman, 1980). If localized spatial frequency filters have an overall two-dimensional gaussian profile, they would optimally specify the combination of spatial frequency and spatial location information, and could be utilized to characterize any complex stimulus. Thus a visual scene could be broken down into a mosaic of, say, 1,000 patches and the total information in the scene could be represented by the outputs of an array of Gabor filters with different frequency, orientation, and phase tuning within each patch. A cross section of such a localized spatial filter is shown in Figure 1.8.

Arrays of Gabor functions (gaussian-tapered sinusoids) are not the only possibility for a patchwise specification of a two-dimensional scene. Young (1986) points out that arrays of gaussian derivatives of different widths in each patch could serve the same purpose. These also have the property of being a complete set, that is of being able to encode any waveform.

Underlying Assumptions

All of our statements thus far about the analysis of complex waveforms into sine wave components are mathematical truisms, but the assumptions underlying them may or may not make them applicable to any given real-life situation. The most critical determinant of whether they can in fact be validly applied to any

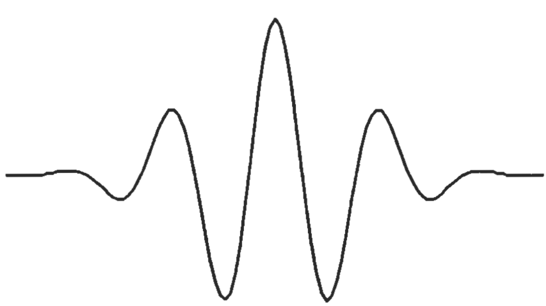

FIG. 1.8 Cross section of a localized spatial frequency filter. This Gabor filter is the product of a sinusoid and a gaussian distribution.

practical situation is the underlying assumption of linearity of summation, or superposition. For instance, suppose that when tested with individual sinusoids a cell gives 10 spikes to a pattern of frequency f and 5 spikes to a pattern of $2f$. Will it give 15 spikes to the combination of $f + 2f$ when they are simultaneously presented? Slight deviations from linearity of summation will only slightly change the result, but clearly any large nonlinearities would render the process of Fourier analysis or Fourier synthesis invalid.

An assumption of linearity is justified in the case of the physics of light, but it is much less obviously so in the case of the physiology of the retina or brain, or in the operation of the visual system as a whole. In fact, it is well known that some visual processes are very nonlinear indeed. For instance, the first quantitative visual relationship historically established—Weber's law—implies a logarithmic rather than a linear relation between light intensity and brightness. A number of nonlinearities have also been shown between certain physical variables and the output of cells in the visual system. It can thus not be assumed automatically that linear analysis can be meaningfully applied to any physiological or psychophysical process. However, most large deviations from linearity occur only under restricted and often rather unusual circumstances. For instance, the large deviations from linearity seen in Weber's Law demonstrations become apparent when brightness is measured over, say, a millionfold range of light intensities. But under ordinary visual circumstances, the range of light intensities within a given scene is usually no more than about 20 to 1. Within such a restricted range, the difference between a linear and a logarithmic function would be slight indeed. In practice, then, the visual system may approximate a linear system to a greater extent than one would assume from classical studies.

A further assumption underlying Fourier analysis is that of homogeneity or spatial invariance, namely, that the underlying structure is everywhere uniform in its properties. This is generally true for any but rather poorly constructed lenses, but it is certainly not so for the retina. As is well known, the central, foveal region is quite different in its properties from the periphery. Foveal cones are smaller and much more densely packed than those in peripheral regions, for instance. So the visual properties of the system are very nonhomogeneous across the whole retina. However, as discussed in the section on global versus local Fourier analysis, we will be concerned exclusively here with quite local properties of the system, rather than of the system as a whole. It is clear that the system itself deals (in the early processing at least) with purely local characteristics of the environment. Over such small distances, the retina is in fact quite homogeneous.

Different Applications of Fourier Analysis

We will be making use of the techniques of Fourier analysis in much of this book. It is important in this regard to differentiate between its application to the study of visual problems and the possibility that the visual system itself may be utilizing such techniques in dealing with visual stimuli. We will be concerned with both of these applications, but it is crucial to distinguish between them.

Linear systems analysis can be (and of course has been) applied to the study of complex waveforms in many physical systems, e.g., voltage waveforms in electronics, light going through optical systems, and seismic waves in the earth. The essence of such an analysis is that of breaking complex waveforms down into separate frequency components, thereby providing a means by which one can quantitatively characterize the shape of the waveform. Insofar as the visual system is approximately linear, a similar application can reasonably be made to quantify neural or psychophysical responses to complex waveforms (patterns) of light. For instance, if it is known (as it is) that the visual system cuts off at a particular high spatial frequency and attenuates low spatial frequencies as well, one can make certain predictions about how a visual scene will be perceived. Some details or overall characteristics of the scene will be invisible, others less prominent, etc.

Quite different from applying Fourier analysis to a quantitative study of some complex waveform such as seismic waves is the construction of a machine (or a computer program, or some other device) to *carry out* a Fourier analysis. For instance, one could build a bank of electronic filters tuned to different temporal frequencies into which a complex voltage waveform could be fed. The various outputs of the individual filters would then indicate the extent to which the input pattern had power in each frequency band, and this system of filters would thus act as a Fourier analyzer. Physiological mechanisms, as well as electronic or mechanical or computational ones, could conceivably be built to carry out such a frequency analysis. In fact, it has long been considered that the cochlea of the ear performs such an analysis of auditory waveforms, the different regions of the basilar membrane being tuned to different temporal frequency bands, so that each cochlear region responds insofar as the input sound wave has power at the frequencies to which that region is tuned. It is conceivable that the visual system is so built as to do a similar frequency analysis of the distribution of light across localized regions of space. An essential requirement for that to occur (and for which we will examine the psychophysical, physiological, and anatomical evidence) is that among the cells processing information from each given region in space there be ones tuned to each of a variety of different spatial frequencies. In the last chapter of the book we have a discussion of possible advantages and disadvantages of the visual system's having apparently evolved a physiological organization at early synaptic levels that approximates a patchwise Fourier analysis of visual space.

SUMMARY

Fourier analysis is a powerful tool for the study of complex waveforms, allowing one to specify quantitatively the characteristics of any complex waveform or shape. An additional major advantage in its application to vision is that it gives a common basis by which one can examine optical, physiological, and psychophysical data. The principal limitation in its application to visual problems is the underlying assumption of linearity, a condition that is only met by the visual system under limited conditions. Finally, it is also possible that the visual system is so constructed as to itself carry out something approximating a local Fourier analysis in its attempt to deal with complex visual patterns.