

# Chapter 2

## Spectral Analysis and Fourier Series



### 2.1 Musical Sounds

As discussed in Chap. 1 in the case of vibrating strings and organ pipes, there are generally many different modes in which a resonant system may vibrate. (See Fig. 1.6 and related discussion.) Generally, more than one of these modes are excited simultaneously in the sounding of a musical instrument. Indeed, their presence or absence is what determines the beauty of a particular tone as well as the difference in sound from one instrument to another. Which modes are excited is not only determined by the characteristics of the resonant system but also by the way in which it is excited. For example, the slipping of the violin string on the bow, or the vibration of the reed in an oboe or a bassoon excites a particular set of modes in those instruments.

Musicians often refer to the extra sounds produced above the pitch of a note as overtones. In most cases, these overtones are harmonically related to the fundamental frequency in that their frequencies are integral multiples of the fundamental. Unfortunately, that fact sometimes produces confusion between the meaning of the musician and that of a scientist analyzing the tone. For example, the first overtone of a vibrating string is actually the second harmonic of the fundamental pitch (its frequency is twice the fundamental frequency), the second overtone is the third harmonic (or three times the normal pitch), and so on.

Life is further complicated by the fact that some instruments produce overtones that are *not* harmonically related to the fundamental. Examples are the kettledrum (or timpani), bells, bars, wood blocks, and so on. Even more surprisingly, the plucked string turns out to have overtones that are not precisely harmonics of the fundamental, whereas the bowed string does.

---

The original version of this chapter was revised: Equations on pages 40 and 41 were corrected. The correction to this chapter is available at [https://doi.org/10.1007/978-3-319-92796-1\\_8](https://doi.org/10.1007/978-3-319-92796-1_8)

In most instruments characterized by harmonically related overtones, the excitation mechanism itself produces a locking effect causing the various harmonics to be in phase with the fundamental. That results from the action of the vibrating reed in woodwinds, the vibrating lips in the case of brass instruments, and the stick-slip motion of the string against the bow in instruments of the violin family.

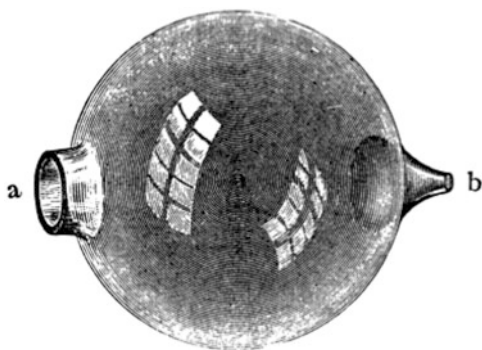
Because the relative distribution of overtones plays such a key role in defining the characteristics of musical sound, it will help to review some of the methods used to determine spectral distributions. By the term “spectral distribution,” we mean the variation of amplitude or intensity of a waveform with frequency. (This chapter gives a qualitative discussion of several important methods of spectral analysis, whose mathematical bases are derived in the appendices.)

## 2.2 Early Methods of Spectral Analysis

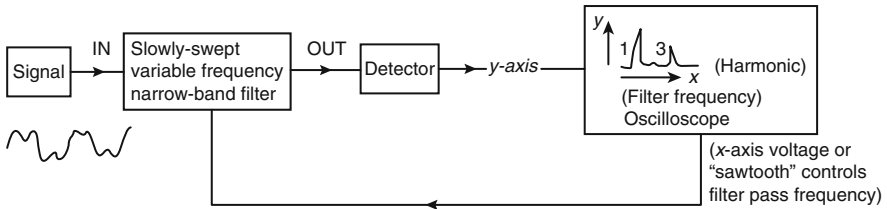
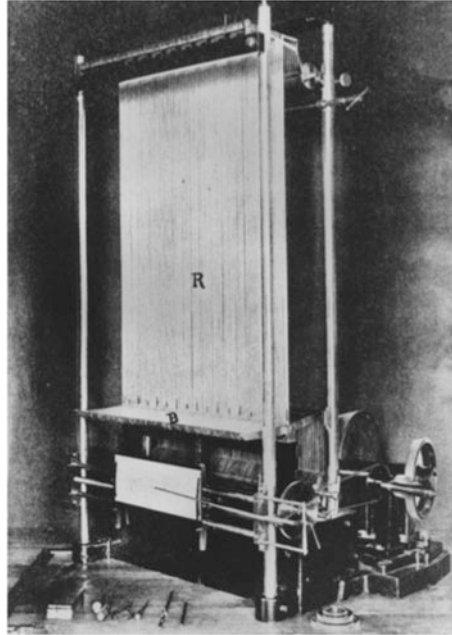
The early pioneer, Helmholtz, did much of his experimental research in acoustics using volume resonators for which the design bears his name. His basic idea was to have a large spherical volume of air resonant at a given frequency that could be driven by sound waves entering through a small aperture. (See Fig. 2.1.) A much smaller tube at the other side of the sphere was designed to fit snugly through a wax seal into his ear so as to block off external sounds. Helmholtz had a set of such matched resonators made that were tuned to different frequencies and managed to accomplish an amazing amount of research with this relatively crude type of apparatus.

Michelson (1903) designed another kind of spectrum analyzer to study the fringes produced in optical interferometry. His analyzer (Fig. 2.2) consisted of a large number of vertical rods tuned to different frequencies. These would vibrate at sympathetic resonances in the audio range when a horizontal lever was made to trace out a particular waveform. The extent of vibration of each rod was recorded on paper and thus provided a measure of the spectral amplitudes.

**Fig. 2.1** An original Helmholtz resonator. Sound entered the resonant volume at *a* and was monitored through the narrow tube at *b*, which was covered with wax molded to fit the experimenter’s ear. From Helmholtz (1885, pp. 43, 373). The resonant frequency is derived in Appendix A. See Eq. (A.68)



**Fig. 2.2** Michelson's spectrum analyzer consisting of vibrating rods tuned to different frequencies. From Michelson (1903, p. 67)

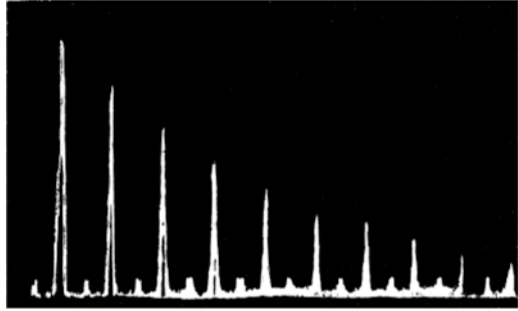


**Fig. 2.3** Schematic diagram of an analog electronic spectrum analyzer. The illustration shows the waveform from a closed organ pipe being broken into its spectral components—principally, the first and third harmonics (in practice, such devices often use one very good narrowband filter at a fixed high frequency which looks at the difference frequency produced when the audio input signal is multiplied by a sine wave from a swept, high-frequency oscillator)

A more-sophisticated electronic approach was developed at the Bell Laboratories during the 1930s to study sound in which a signal could be recorded on a magnetic disc that was repeatedly scanned while a narrowband filter swept slowly through the audio frequency range. The spectra were shown by using the rectified output of the filter to darken a piece of paper. Apart from the time required to observe the spectrum, the recording medium had very limited dynamic range.

A more recent version of that approach having a much wider dynamic range is shown in Fig. 2.3. The main difficulty with such analyzers is that you need a sustained source of sound (or a tape loop) for analysis (Fig. 2.4).

**Fig. 2.4** Odd-harmonic spectrum from a square wave determined with this analyzer



## 2.3 The Decibel (dB)

Relative spectral amplitudes are often described in terms of “dB,” or decibels. The “decibel” was originally called the “transmission unit” and referred to the loss in a standard length of telephone wire (Martin 1924). It was subsequently renamed in honor of Alexander Graham Bell, but with his name misspelled and entered in lower case. The more recent abbreviation for the unit (the “dB”) at least capitalizes the “B.” The most important thing to remember about the unit is that it represents a logarithmic measure of the *ratio* of two intensity (or power) levels. Specifically, the ratio of the intensity levels  $I_2$  to  $I_1$  is defined in dB as

$$10 \log_{10} \left( \frac{I_2}{I_1} \right). \quad (2.1)$$

Because the intensity is generally proportional to the square of an amplitude (for example,  $I_2 \propto A_2^2$  and  $I_1 \propto A_1^2$ ), the intensity ratios in dB may also be written as<sup>1</sup>

$$10 \log_{10} \left( \frac{A_2^2}{A_1^2} \right) = 10 \log_{10} (A_2/A_1)^2 = 20 \log_{10} (A_2/A_1) \quad (2.2)$$

where  $A_2$  to  $A_1$  is the amplitude ratio. If the wave is attenuated in passing through a medium, the result is a negative number of dB, and vice versa. Some useful benchmarks to keep in mind: 10 dB corresponds to an intensity ratio of 10:1 whereas doubling the intensity only amounts to about 3 dB. On the other hand, doubling the amplitude results in a gain of about 6 dB. Conductor Leopold Stokowski became enamored with decibels during the 1930s. People at the Bell Laboratories gave him a dB meter hooked to a microphone which he used on the podium of the Philadelphia Orchestra. One can imagine comments during rehearsals such as, “Mr. Tabuteau, I’d like 6 dB more in the crescendo at letter A.”

<sup>1</sup>Note by definition,

$a = \log_{10} b$  means that  $b = 10^a$ .

Therefore,

$\log_{10} b^2 = \log_{10} 10^{2a} = 2a = 2 \log_{10} b$

**Table 2.1** Sound pressure levels (SPLs) referred to  $2 \times 10^{-4}$  dynes/cm<sup>2</sup> (100 dB =  $1 \mu\text{W}/\text{cm}^2$ )

SPL (in dB)	Source
200	16 in. naval gun at 12 ft
140	Jet taking off
130	Jet taxiing
125	Student rock concert at Yale university at 50 ft
120	Threshold of pain
110	Construction site (pneumatic drills at 100 ft)
100	Lawn mower
95	Good Stradivari violin at 3 ft; most car interiors; times square traffic.
90	San Francisco symphony in Carnegie hall playing Mahler (100 ft)
85	Kirov orchestra in Verizon hall playing Wagner (100 ft)
80	City street; alarm clock (at 2 ft)
75	Frappuccino maker at 3 ft
70	Shouting at 4 ft
60	Normal conversation at 3 ft; busy office
50	Quiet office or classroom
40	Living room
30	Bedroom at night
20	Recording studio
10	Yale maintenance workers
0	Snowflake hitting ground

Note: These are total SPLs and not “A weightings.” Extended listening to sound levels above 90 dB is thought by OSHA to be damaging to the ear. The Bavarian Radio Symphony Orchestra was cited for violating a new law on noise level while rehearsing the “State of Siege” by Dror Feiler, a piece containing sustained sound levels of 97 dB (N.Y. Times, April 20, 2008, p. 1)

Confusion may be introduced by people who refer to absolute sound levels in decibels. What they usually mean by that terminology is that the sound intensity ratio is in respect to a standard reference level where 0 dB corresponds to  $2 \times 10^{-4}$  dynes/cm<sup>2</sup>. That value is approximately the threshold of hearing at 2 kHz. On that same scale, 120 dB is about the threshold of pain. By coincidence, an increment of 1 dB is about the smallest change in intensity ratio that the average human ear can detect, although the value varies somewhat with individuals, with frequency, and with sound level. (See Riesz 1932.) Using such SPL (“Sound Pressure Level”) meters, the various peak absolute sound levels shown in Table 2.1 were obtained.

Recently, circuits containing multiple frequency-band transmission filters to cover the audio spectrum have flooded the “Hi-Fi” market and are generally calibrated in decibels. The frequency bands are often spaced at octave, or even one-third octave, intervals and use light-emitting diodes to indicate the relative sound intensity levels within the different bands. Although they have the advantage of rapid response and can provide a rough portrayal of spectra in “real time,” the resolution is limited by the number of filters one can crowd into a small circuit. A nice application of this display has been incorporated in the sound level meter

**Fig. 2.5** A portable real-time spectrum analyzer and sound level meter



made by the AudioSource company. As illustrated in Fig. 2.5, the meter is portable and provides a real-time spectral display of absolute sound levels in dB detected from a calibrated condenser microphone.

With the computer methods discussed below, one can increase the resolution merely by increasing the dimension of a column array. It used to be that machine “running times” were an impediment to mathematical analysis. But, now that we have computer speeds in the GHz domain and nearly unlimited random access memory storage capability, Fourier analysis can be done throughout the entire audio spectrum in real time with good resolution.

## 2.4 Fourier Analysis

Although the mathematical techniques involved in Fourier analysis have been known since the 1820s, the ability to use this method rapidly in real-time analysis required innovations in computer technology that did not arise until the 1970s. One major advantage is that you do not need a very long sample of a waveform to determine its spectral distribution. Indeed, if you know that the waveform is truly periodic, you only need one period for analysis. Hence, in many cases, the spectra can be captured “on the fly.”

## 2.5 A Brief Historical Background of Fourier Series

Fourier began the mathematical work which led to his formulation of what we now call “Fourier Series” in a theoretical study of heat flow in 1807, stimulated by engineering problems encountered in the boring of Napoleon’s cannons.<sup>2</sup> Fourier solved the heat-flow equation for sinusoidal distributions of temperature. But, he needed an infinite series of such solutions to describe the results of arbitrary temperature distributions on the walls of the material. Fourier’s initial paper on this subject was highly controversial. Several outstanding mathematicians did not believe what he was saying and urged the paper’s rejection. In fact, his paper was refused publication until Fourier himself was elected President of the French Mathematical Society in 1822.

According to Whittaker and Watson (1920), the major background developments were as follows:

1. *D’Alembert* had solved the wave equation for the vibrating string problem and obtained a solution of the form

$$y(x, t) = 1/2[f(x + ct) + f(x - ct)]. \quad (2.3)$$

Note that  $y = f(x)$  is the shape of the string at time  $t = 0$ .

2. *Daniel Bernoulli* next showed that a formal solution to the problem was also given by a sum of solutions of the type summarized in Chap. 1 by Eq. (1.13):

$$y(x, t) = \sum_{n=1}^{\infty} (A_n \sin(n\pi x/L) \cos(n\pi ct/L)) \quad (2.4)$$

where the  $A_n$  are adjustable constants.<sup>3</sup> Bernoulli went on further to claim that this result was the most general solution to the problem possible. (Although the claim sounded like a Madison Avenue advertising slogan, it turned out to be right.)

3. Neither *d’Alembert* nor *Euler* believed Bernoulli and protested that such a series could not possibly converge to a function such as  $f(x) = x(L - x)$  at

---

<sup>2</sup>Baron Jean-Baptiste-Joseph von Fourier (1768–1830) accompanied Napoleon in 1798 on his expedition to Egypt, where he served as Secretary for Napoleon’s newly formed Institut d’Egypt. In Cairo, he did extensive research on Egyptian antiquities and gave advice on engineering matters. He returned to France in 1801, about the same time that the Rosetta Stone and other major ancient Egyptian relics were surrendered to the British. Back in France, he was charged with the publication of an enormous mass of Egyptian material which became known as *Description de l’Egypt* (in 21 volumes from 1808 to 1825). He was also the first to describe the atmosphere’s trapping of heat as “The Greenhouse Effect” in the 1820s. (See Segré 2002, p. 119.)

<sup>3</sup>The notation  $\sum_{n=1}^{\infty} (A_n \sin(n\pi x/L) \cos(n\pi ct/L))$  means that you sum the expression for the values of  $n = 1, 2, 3, \dots$  to infinity.

$t = 0$ , or even worse, the boundary conditions at  $t = 0$  on a plucked harpsichord string.

4. *Fourier* (1822) proved for the first time that such a series did indeed converge in a large number of specific cases while discussing his analytic theory of heat flow.
5. Others (*Poisson*, *Cauchy*, *Dirichlet*, and *Bonnet*) went on to attempt more general proofs (some of them wrong). According to Whittaker and Watson (1902), the first correct proof of convergence was given by Dirichlet.

## 2.6 A Note on the Convergence of Infinite Series

The concept of convergence of a sum such as that in Eq.(2.4) at  $t = 0$  is of fundamental importance in establishing the usefulness of Fourier series. For a rigorous discussion of convergence, the reader should consult a treatise on mathematical analysis such as that by Whittaker and Watson (1902). What follows here is a more pragmatic approach to the problem.

Suppose we have a sum of numbers of the form

$$S = a_1 + a_2 + a_3 + \cdots + a_n \cdots$$

where the  $n$ th term is a known function of  $n$ . For the sum to converge to a limiting value,  $a_n$  clearly must go to zero as  $n \rightarrow \infty$ . Although that is a necessary condition for convergence, it is not a sufficient one. For example, the well-known series

$$S = 1 + 1/2 + 1/3 + \cdots + 1/n + \cdots$$

does not converge, but obviously satisfies that “necessary” condition. Convergence does occur when  $a_n + 1/a_n$  goes to zero in the limit that  $n \rightarrow \infty$ . (The divergent case quoted above obviously does not satisfy that requirement.)

In the present computer age, it is often adequate to run off the sum of the series to a few dozen terms to see what actually happens. In that approach, if you stop calculating the sum after  $|a_n| < 10^{-7}|S|$ , you will usually have reached the convergence limit within the accuracy of the computer. That is, “single-precision” computer calculations in which the mantissa is evaluated to 24-bit accuracy are typically good to only about one part in  $10^7$ . (Of course, convergent series can always be computed in extended precision using more bits for the calculation.)

A numerical example will help for clarification. The infinite series for  $e^x$  is given by

$$S = 1 + x + x^2/2 + x^2/3! + \cdots + x^n/n + \cdots \quad (2.5)$$

It is useful to note that the  $n$ th term of the series is easily related to the  $(n - 1)$ th term by



**Fig. 2.6** The first 30 terms for the series in Eq. (2.5) for  $x = 10$

n	a (n)	s
1	0	1
2	10	11
3	50	6.1
4	166.667	22.77
5	416.667	64.44
6	833.334	147.77
7	1388.889	286.66
8	1984.127	485.07
9	2480.159	733.09
10	2755.732	1008.66
11	2755.732	1284.24
12	2505.211	1534.76
13	2087.676	1743.52
14	1605.905	1904.11
15	1147.075	2018.82
16	764.717	2095.29
17	477.948	2143.09
18	281.146	2171.2
19	156.193	2186.82
20	82.207	2195.04
21	41.104	2199.15
22	19.573	2201.11
23	8.897	2202
24	3.869	2202.39
25	1.612	2202.55
26	0.645	2202.61
27	0.248	2202.64
28	0.092	2202.65
29	0.033	2202.65
30	0.012	2202.65

$$a_n = a_{n-1}x/n . \tag{2.6}$$

The series will converge for any finite value of  $x$  because

$$a_n/a_{n-1} \rightarrow 0 \text{ as } n \rightarrow \infty .$$

The first 30 terms for the series are illustrated in Fig. 2.6 for the case  $x = 10$ . As can be seen from the figure, the increment  $a_n$  rapidly builds up for the first few powers of  $x$  but goes through a maximum value at about the 11th term. After that, the  $n!$  in the denominator rapidly reduces the increment to zero and the series converges to 1 part per million by the 30th term to  $S = 2202.65$ . To get the numerical value for  $e$  ( $= 2.718282\dots$ , the base of the Naperian logarithms), one merely lets  $x = 1$  in the series Eq. (2.5). The number  $\pi$  is also the result of a convergent infinite series, as are all the transcendental trigonometric functions.<sup>4</sup>

---

<sup>4</sup>Ramanujan (1914) gave the most rapidly convergent series for  $1/\pi$  ever discovered:  $\frac{1}{\pi} = \frac{1}{4} \left[ \frac{1123}{882} - \frac{22,583}{882^3} \cdot \frac{1}{2} \cdot \frac{1 \cdot 3}{4^2} + \frac{44,043}{882^5} \cdot \frac{1 \cdot 3}{2 \cdot 4} \cdot \frac{1 \cdot 3 \cdot 5 \cdot 7}{4^2 \cdot 8^2} - \dots \right]$ . Amazingly, the first term by itself gives  $\pi = 3.141585041\dots$  (Ramanujan liked to entertain his friends by reciting the endless digits of  $\pi$  at parties.

### 2.7 Specific Examples of Convergence for Periodic Series

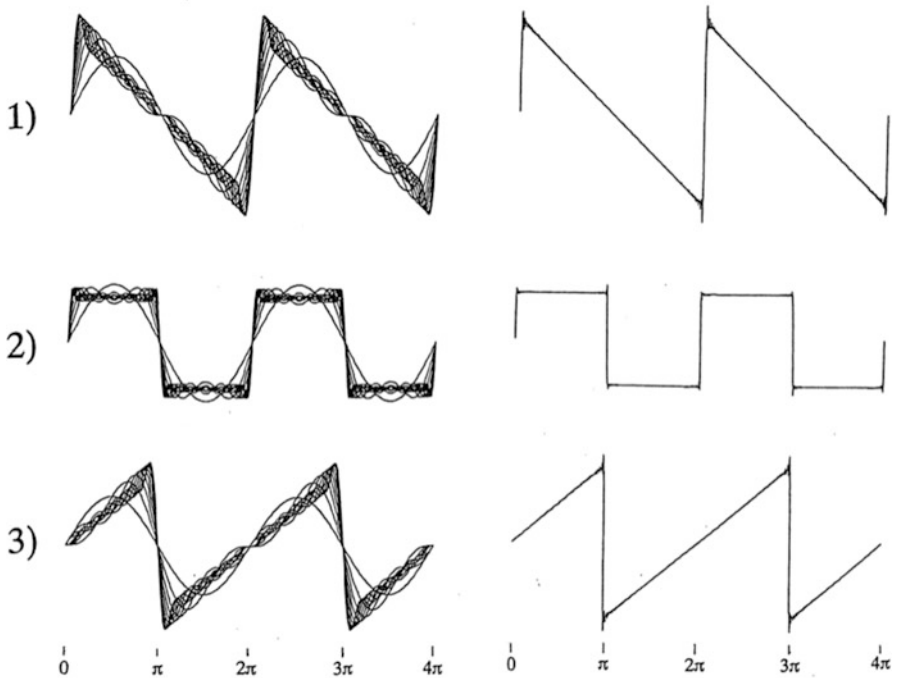
The following three examples involve convergence of an infinite series for each value of  $x$  over the domain  $0 \leq x \leq 4\pi$ . All three represent periodic functions that repeat themselves over the range from 0 to  $2\pi$ . (The range for  $x$  from 0 to  $4\pi$  was chosen to illustrate two periods of the function in each case.) Here, we have used a computer to demonstrate convergence by adding up the terms for different values of  $n$  at each value of  $x$ . For each of the three cases listed below, a superposition of the first ten terms is shown at the left in Fig. 2.7, and the limit of the series after 100 terms is shown at the right. Although the three cases look superficially similar, the results converge in each case to very different, highly non-sinusoidal functions.

Case 1: "Sawtooth":

$$y = \sin x + \frac{1}{2} \sin 2x + \frac{1}{3} \sin 3x + \frac{1}{4} \sin 4x + \dots + \frac{1}{n} \sin nx \tag{2.7}$$

Case 2: "Square Wave":

$$y = \sin x + \frac{1}{3} \sin 3x + \frac{1}{5} \sin 5x + \dots + \frac{1}{n} \sin nx \text{ [} n \text{ odd]} \tag{2.8}$$



**Fig. 2.7** Convergence of the three infinite series shown in the text over two fundamental cycles ( $0 \leq x \leq 4\pi$ .) Case (1) Sawtooth waveform; (2) Square wave; and (3) the Gibbs Zigzag. The column on the left shows the superposition of the buildup of the series through the first 10 terms. The column on the right shows the series after 100 terms were added in each case

Case 3: Gibbs “Zigzag”:

$$y = \sin x - \frac{1}{2} \sin 2x + \frac{1}{3} \sin 3x - \frac{1}{4} \sin 4x + \cdots + \frac{(-1)^n}{n} \sin nx \quad (2.9)$$

## 2.8 The “Gibbs Phenomenon” or Wilbraham Effect

If you look in the vicinity of the vertical discontinuities in the development of the infinite series shown in Fig. 2.7, you will notice a small “horn” sticking up above the waveform. That effect was first discovered by the Scottish mathematician Henry Wilbraham in 1848. It was rediscovered some 50 years later by Gibbs (1899) and has since come to be known as the “Gibbs Phenomenon” in Fourier Series. The width of the horn gets narrower and narrower as the number of terms added to the series increases, but it never disappears. It arises because the convergent limit of the series at the discontinuity differs by about 14% from the limit a small distance away on the curve. Since the waveforms of interest in the present study are not characterized by vertical discontinuities, the effect does not show up in musical instrument waveforms and is merely of historical interest here.

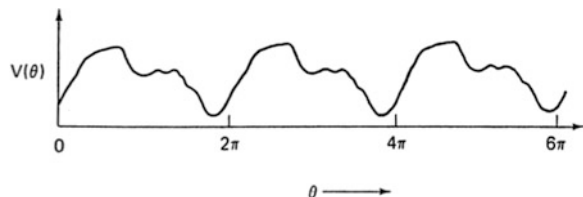
## 2.9 Basic Aspects of Fourier Series

In what follows here, we will restrict ourselves to periodic functions that are “well-behaved” in the sense that they are continuous and their slopes are finite. By a periodic function  $V(\theta)$  such as shown in Fig. 2.8, we mean that

$$V(\theta + 2\pi) = V(\theta). \quad (2.10)$$

Many musical instrument waveforms are periodic in the time, or at least quasi-periodic after an initial excitation transient has died down. For example, the sound pressure wave produced by a closed organ pipe is shown in Fig. 2.9, where the pipe was turned on at the start of the oscillogram. As can readily be seen by eye, the waveform settles down to a periodic one after about ten cycles of the fundamental pipe resonance.

**Fig. 2.8** A hypothetical periodic waveform



**Fig. 2.9** Oscilloscope display of the waveform from a quintadena (closed organ pipe of circular cross-section)



As Fourier showed, any such periodic function can be represented by an infinite series of harmonics of sine and cosine functions over the fundamental period. Thus,  $V(\theta)$  in Eq. (2.10) could be written

$$V(\theta) = C_0 + A_1 \sin 1\theta + A_2 \sin 2\theta + A_3 \sin 3\theta + \dots + B_1 \cos 1\theta + B_2 \cos 2\theta + B_3 \cos 3\theta + \dots \quad (2.11)$$

or

$$V(\theta) = C_0 + \sum_{n=1}^{\infty} A_n \sin n\theta + \sum_{n=1}^{\infty} B_n \cos n\theta \quad (2.12)$$

Here, the constant  $C_0$  allows for a net DC (“Direct Current”) offset of the waveform from the horizontal axis. The terms involving  $\sin n\theta$  and  $\cos n\theta$  are called the  $n$ th harmonic terms and the coefficients  $A_n$  and  $B_n$  are called harmonic amplitudes.

## 2.10 Calculating the Fourier Coefficients for $V(\theta)$

We could just postulate different values for the coefficients  $C_0$ ,  $A_n$ , and  $B_n$ , and evaluate the series in Eq. (2.11) with a computer in the same way that we computed those for Fig. 2.7. We might even try to narrow in on a set of coefficients that would match a particular waveform. However, that would be an extremely tedious and inefficient approach. Fortunately, Fourier worked out a systematic method to compute the coefficients from the waveform directly. The method involves integral calculus and is described in detail in Appendix C. Essentially, the different coefficients are determined by finding the areas under various curves related to the initial waveform over one fundamental period. The DC constant  $C_0$  is the average value and is determined from the area under the curve for  $V(\theta)$  itself, whereas the coefficients  $A_n$  are determined from the area under the curve,  $V(\theta) \sin(n\theta)$ , and those for  $B_n$  from the curve,  $V(\theta) \cos(n\theta)$ . For musical instruments, the waveforms can be measured numerically using an A-to-D converter, a circuit that converts Analog microphone voltages to Digital output values to be read by a computer. (Microphone voltages are usually proportional to the sound wave pressure.)

Once numerical values have been determined for the sine and cosine terms ( $A_n$  and  $B_n$ ) in the series, it is desirable to express the results in terms of one net coefficient and phase for each harmonic (value of  $n$ ). That process just involves a little trigonometry. We rewrite the original Fourier series in Eq. (2.11)

$$V(\theta) = C_0 + \sum_{n=1}^{\infty} A_n \sin n\theta + \sum_{n=1}^{\infty} B_n \cos n\theta$$

as an equivalent series involving one sine and a phase angle for each harmonic:

$$V(\theta) = C_0 + \sum_{n=1}^{\infty} C_n \sin n\theta + \phi_n. \quad (2.13)$$

One then evaluates the coefficients  $C_n$  and the phases  $\phi_n$  in terms of  $A_n$  and  $B_n$  by comparing like terms in the two different expressions for the infinite series.

Thus,<sup>5</sup>

$$C_n = \sqrt{A_n^2 + B_n^2} \text{ for } n \geq 1 \quad (2.14)$$

and

$$\phi_n = \arctan(B_n/A_n).$$

Often, one is primarily interested in the relative distribution of the net harmonic amplitudes  $C_n$  because they correspond roughly to the psychological impression the sound makes on the human ear. As shown in Appendix C, the relative energy distribution in the harmonics of a Fourier series goes as the square of the amplitudes. Some people prefer to convert that number into decibels because the ear responds logarithmically to the harmonic intensity.

Although it is straightforward (but tedious) to do a Fourier analysis by hand, the calculation is a simple matter with a high-speed computer. A program for doing that is given in Appendix C, together with a derivation of the mathematical quantities involved. Not only is that approach to the problem much faster than the older methods of spectral analysis, you only need one period of the waveform in order to determine the harmonic structure. Thus, you can catch the spectral distribution in the time of one period of the waveform rather than, for example, spending a long time scanning the output of a tape loop (as in Fig. 2.3) while a narrow frequency filter is slowly swept through the spectrum.

## 2.11 An Example of Discrete Fourier Analysis

In order to do the computations involved in Fourier analysis, one needs to sample at least one period of the waveform digitally and obtain lots of points. In doing that sort of analysis myself, I used a high-quality condenser microphone to pick up the sound and fed its output into a high-speed A-to-D (“Analog-to-Digital”) converter controlled by a computer. The computer recorded the data, showed the waveform, did a Fourier analysis, and then displayed the relative amplitudes of the harmonic coefficients. The photograph in Fig. 2.10 was taken during a lecture I once gave at Yale in which a student (William C. Campbell) blew a note on a 50-ft garden hose. The hose behaved like a narrow-scale open pipe with modes spaced at about 11 Hz. Campbell was able to phase-lock a large number of those modes in the mid-audio range (at a fundamental frequency of about 307.7 Hz) and produce a waveform with a sharp, periodic pulse that sounded much like a braying elephant. (The waveform

---

<sup>5</sup>These results are obtained by applying the trigonometry identities  $C_n \sin(n\theta + \phi_n) = C_n \sin n\theta \cos \phi_n + C_n \cos n\theta \sin \phi_n$

and

$$\tan \phi_n = \frac{\sin \phi_n}{\cos \phi_n} \text{ together with } \cos^2 \phi_n + \sin^2 \phi_n = 1.$$



**Fig. 2.10** Photograph taken in Davies auditorium at Yale University during a lecture by the author on “Live Fourier Analysis” in the 1970s. The garden hose was played by Yale student, William C. Campbell

and amplitude spectrum are shown on the oscilloscope in Fig. 2.10, together with the HP-2116B computer used.)

The Campbell waveform provides a nice example of the way in which a sum of sine waves can add up to produce a sharp pulse. At the same time, it provides a useful example to illustrate the convergence of a Fourier series. The amplitude coefficients and phases shown in Fig. 2.11 were computed from the digitized waveform using the program described in Appendix C. A histogram of the Fourier coefficients  $C(n)$  is shown as a function of harmonic number starting from the left with  $n = 1$  in Fig. 2.12, together with the waveform over one cycle. (The DC offset,  $C(0)$ , probably resulted from air coming out of the hose near the microphone and is not included in the histogram.)

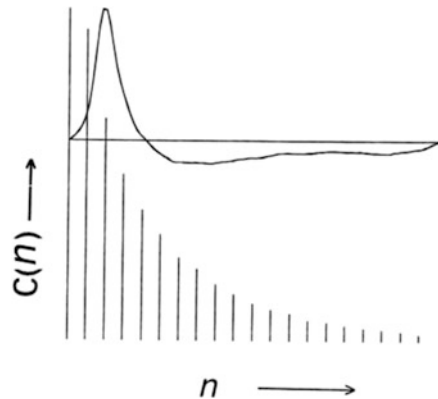
The original waveform can, of course, be reconstructed by putting the amplitudes and phases from Fig. 2.11 back into Eq. (2.12). That process illustrates the convergence of the Fourier series with increasing number of harmonics, as has been done in Fig. 2.13 where the numbers represent the maximum number of harmonics used in the reconstruction. The values of the phase are very important in determining the visual shape of the waveform, whereas the harmonic amplitudes are more related to the sound heard by ear (Fig. 2.14).

The oscillogram of the closed pipe waveform in Fig. 2.9 provides an example of the pitfalls involved in Fourier analysis. If you had started analyzing that data when the organ pipe was initially turned on, you probably would not even have been able to determine the fundamental period. There was an initial transient during which only higher modes of the pipe were excited. Then, as time went on, the fundamental mode slowly built up and became the dominant source of sound in the spectrum. At the extreme right end of the oscillogram, the waveform has become strongly

**Fig. 2.11** Relative amplitude coefficients  $C(N)$  and phases  $P(N)$  computed from one cycle of the garden hose waveform for the first 20 harmonics

N	C(N)	P(N)
0	-5.1666666667	---
1	200.18752265	1.1031844815
2	187.226833595	0.2328547785
3	133.673566323	-0.4481088423
4	99.610983131	-1.0783248202
5	78.33444076	4.5807732471
6	63.36763511	4.0453575124
7	49.62334311	3.4573178766
8	42.773894111	2.9106885433
9	33.244534907	2.241522386
10	27.3404696372	1.7578400142
11	21.7070327263	1.2408095503
12	18.0638132271	0.7281888317
13	15.7394445738	0.1343847924
14	11.7314319947	-0.3948992557
15	10.5061603374	-0.9722905303
16	8.7580766154	-1.5573037672
17	7.1367499029	4.1680712601
18	6.0349917008	3.5319263251
19	4.8830470689	2.8393246777
20	3.2039572595	2.2241875619

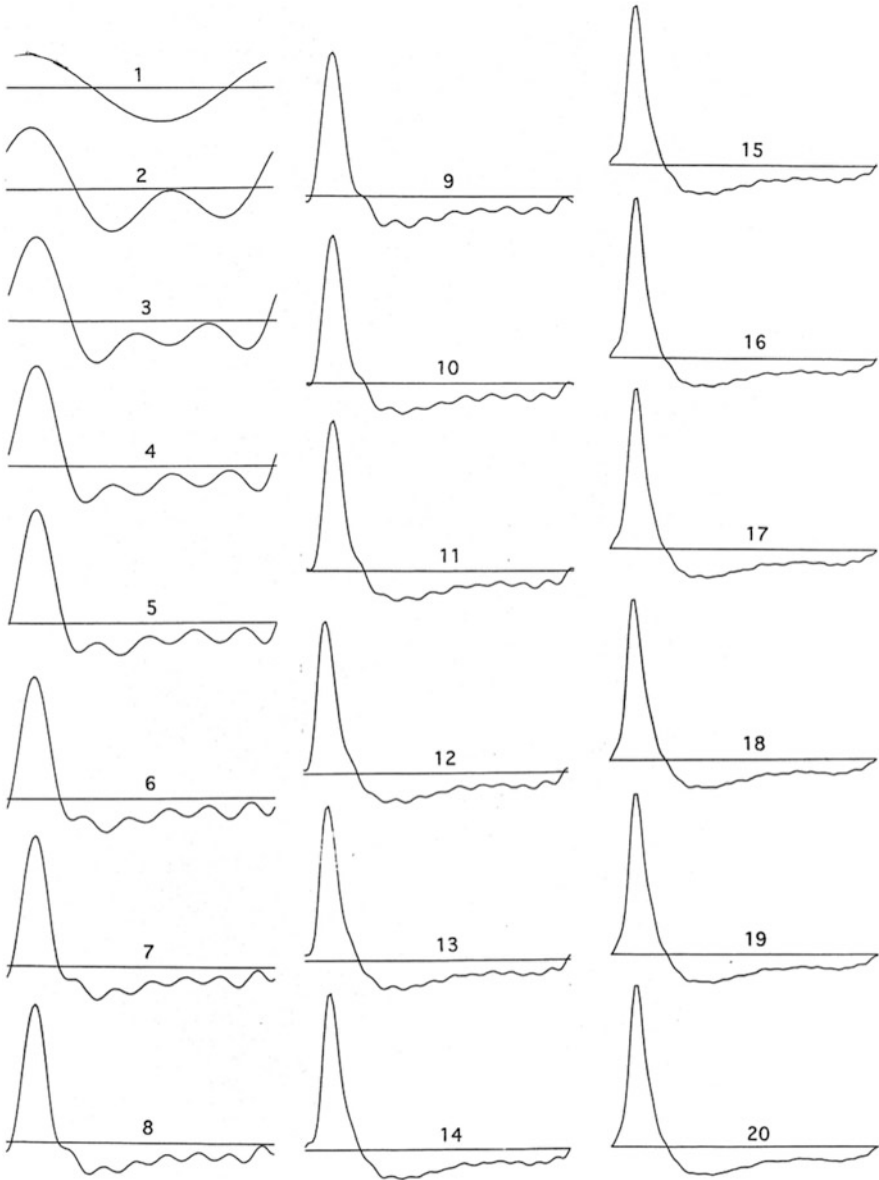
**Fig. 2.12** Waveform and histogram of the Fourier coefficients  $C(n)$  for the garden hose waveform



periodic and one can easily pick out the period. Just by looking at it, you can see that only odd harmonics are of importance in the steady-state waveform and not much more than the first three are significant. If you were to use a slow computer, it would help to estimate just how many harmonics you need to analyze in advance, for the running time in the computation of the discrete Fourier analysis goes up as the square of the number of harmonics (and of the number of points). Once you know the fundamental period, you could, of course, go back to the beginning and Fourier analyze the entire spectrum, period by period. That process would show how the harmonics changed during the transient.

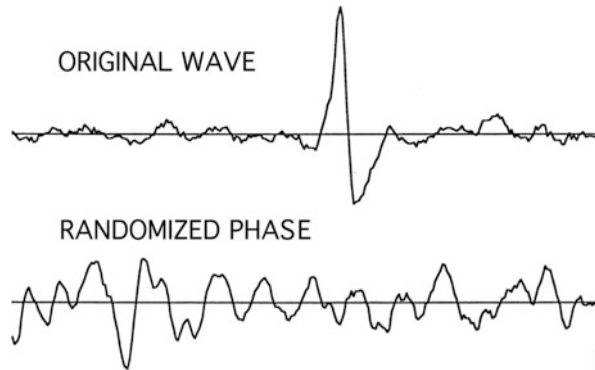
We can use the reconstruction of the waveform from the harmonic coefficients to show how the Fourier series itself converges. That has been illustrated in Fig. 2.13 for the garden hose waveform in Fig. 2.12.





**Fig. 2.13** Reconstruction of the waveform in Fig. 2.12 from the Fourier coefficients. The numbers represent the maximum number of harmonics used in each case to reconstruct the waveform

**Fig. 2.14** Top: Waveform from a tuba at 33 Hz reconstructed from the original 50 amplitudes and phases obtained from Fourier analysis. Bottom: Waveform reconstructed from exactly the same set of 50 amplitude coefficients but with randomly selected phases for the different harmonics



The cochlea in the human ear acts somewhat like a spectrum analyzer in that thousands of different channels respond to sound waves of different frequency and transmit pulses to the brain such that their rate increases with the loudness detected by each channel. To a large extent, the apparent tonal color of the sound is determined by that distribution—hence by the energy content in each harmonic component. The harmonic distribution thus gives the listener the main perception of tonal color. However, that is not the entire story. The ear is also somewhat sensitive to the actual shape of the waveform. Hence, a waveform consisting of periodic sharp spikes sounds somewhat different from that produced by a periodic waveform with the same relative harmonic amplitudes but different phases. The relative phases in a musical instrument waveform are usually determined by the excitation process—for example, the stick-slip mechanism in the bowed violin string, the vibration of the reed in a woodwind, or that in the lips of a brass instrument player. Generally, these processes produce phase locking of the different harmonics in respect to the fundamental period of the instrument so that the phases do not just wander around randomly.

A question naturally arises regarding the number of points needed for analysis of the waveform. Most instruments seldom have more than 10–30 important harmonics. (There are exceptions such as the low notes on a krummhorn, or tuba.) It is, of course, the relative distribution of the stronger harmonics that mainly determines the tonal color (not to mention their variation with time, as in the case of vibrato). Surprisingly, a criterion developed many years ago by Harry Nyquist (1924) for the transmission of telegraph pulses is relevant. He showed in general that in order to transmit signals digitally, one needs to sample the original analog signal at more than twice the maximum frequency you want to transmit. His criterion (following from something called the “Nyquist Sampling Theorem” and which plays a key role in the CD-recording industry) also works backwards. For example, if you have a waveform whose fundamental frequency is 200 Hz and want to examine 20 harmonics (i.e., up to a frequency of 4 kHz), you need to have more than 8000 samples per second, which means more than 40 points over one period of the waveform. In practice, you would want to exceed that minimum

requirement by at least a factor of 2 or 3, which means that you would probably want at least 100 points over the fundamental cycle. (The minimum number required by the Nyquist criterion would just give two samples over the period of the highest frequency component.) However, for similar reasons, it is also pointless to analyze the waveform for a number of harmonics larger than half the number of points measured over one period. (If you do so, you just get the same spectral information back again in the higher harmonics, but in reverse order.) It is sometimes implied that there are only slight differences in the waveforms between different instruments and that a relatively small fraction of the sound intensity falls in the overtones. Nothing could be farther from the truth. The differences in the harmonic structure can be enormous, even between instruments of the same species. We will illustrate that fact with a variety of different examples throughout this book.

## 2.12 The Fourier Transform

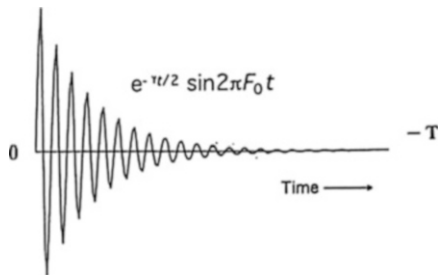
The method of Fourier series discussed above works well when one has a precisely periodic waveform. However, in practice one often encounters situations where the waveform may be quasi-periodic, but varies significantly over the time of observation. Such cases might include the sound from an instrument played with vibrato, or the sound from an instrument such as a harpsichord or piano that is inherently transient in character. Finally, there are some instruments (e.g., tympani and bells) where the waveforms are not even approximately periodic and for which the overtones are not harmonically related. Here, there is a useful computational method based on Fourier analysis that goes under the heading Fourier Transform. In that approach, we observe the wave over a very long time  $T$  that is not the period of the vibration as discussed before. We then pretend that the waveform is periodic over that long time interval  $T$  (which generally includes many cycles of oscillation in the frequency range of interest). Now, even though it is just a mathematical fiction, we can apply our previous results for periodic waveforms to compute the harmonics present with fundamental frequency  $1/T$ . But, of course, the results will not apply outside the region  $0 \leq t \leq T$ . Most of the spectral components will be of no physical interest. However, the components we do care about will be contained in the computed frequency range and appear as harmonics of  $1/T$ .

It will help to illustrate with a particular example. Consider a decaying waveform of the type

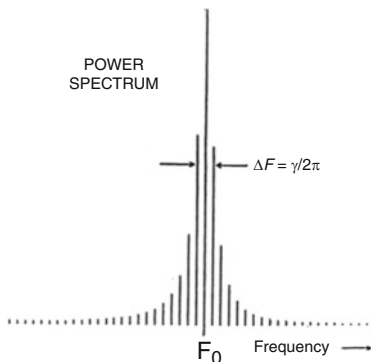
$$y(t) = \exp(-\gamma t/2) \sin(2\pi F_0 t) \text{ for } 0 \leq t \leq T \quad (2.15)$$

which might represent the sound amplitude from the fundamental mode of a string plucked at  $t = 0$ . (See Fig. 2.12.) Here,  $\gamma/2$  is the amplitude decay rate which might result from energy being coupled to a sounding board. Because the energy in the wave motion varies as the square of the amplitude, the energy decay rate in this situation is simply  $\gamma$ , or twice that for the amplitude decay rate. We assume here that

**Fig. 2.15** Decay of a damped waveform given by Eq. (2.14)



**Fig. 2.16** Energy spectrum for the waveform in Eq. (1.14) computed with a Discrete Fourier Transform



$T \gg 1/F_0$ , or equivalently, that Eq. (2.14) describes the oscillation of the string over many fundamental periods of the string oscillation frequency. In practice, there might be a thousand or more digital samples taken of the string amplitude during the long time interval  $T$ .

Fourier analysis of the waveform described by Eq. (2.14) and Fig. 2.15 results in the spectrum shown in Fig. 2.16, where the square of the net Fourier amplitudes is plotted as a histogram as a function of the harmonic number. Figure 2.15 represents the energy distribution in the spectrum. Note that the spectrum peaks at  $F_0$ , which itself is a high harmonic of  $1/T$ , where  $T$  again is the long observation time. As shown in Appendix A, the energy or power distribution in this case has a so-called “Lorentzian shape” with a full width at half-maximum intensity of  $\Delta F = \gamma/2\pi$ . That is, the resonance is not perfectly sharp but is spread over a range of frequencies  $\Delta F$  centered about  $F_0$ . The spread arises because the initial signal is not a constant single-frequency wave persisting for an infinite length of time. The result is actually an example of the Uncertainty Principle—something known to electrical engineers long before Heisenberg made his famous pronouncement as applied to quantum physics. As we have shown here, it is a consequence of Fourier analysis that

$$\Delta F \Delta t \approx \frac{1}{2\pi} \tag{2.16}$$

where  $\Delta t \approx 1/\gamma$ . To put it in different words, the limiting uncertainty in the frequency measurement ( $\Delta F$ ) varies approximately as the inverse of the signal duration ( $\Delta t$ ).<sup>6</sup>

## 2.13 Window Functions

Although the Discrete Fourier Transform worked perfectly well with the waveform shown in Eq. (2.14), there are some pitfalls in the method. The waveform in Eq. (2.14) was carefully chosen to go to zero at  $t = 0$  and to become very small as  $t \rightarrow T$ . However, an arbitrarily chosen wave shape,  $y(t)$ , might be nonzero at both  $t = 0$  and  $t = T$  and result in a function that could not conceivably be periodic in the large time interval without having major discontinuities. They, in turn, would result in spurious frequency components during Fourier analysis. To avoid that difficulty, it has become a standard practice to multiply the data obtained in the large time interval by a Window Function which we will call  $W(t)$  that goes smoothly to zero at both  $t = 0$  and  $t = T$ . Although the process tends to broaden the computed spectral widths and produces minor distortion of resonant line shapes, it does not interfere with the determination of the resonant frequencies and, most important of all, it does not introduce spurious spectral components. There are almost as many window functions as people who have worked in this field. The most commonly used one is that proposed initially by the Austrian mathematician, Julius von Hann, which for some strange reason is now called the ‘‘Hanning Window.’’<sup>7</sup> It multiplies the data by the function

$$W(t) = 0.5[1 - \cos 2\pi t/T] \text{ for } 0 \leq t \leq T. \quad (2.17)$$

This ‘‘Hanning Window’’ has been adopted as a standard by the IEEE (‘‘Institute for Electrical and Electronics Engineers’’) and is built into a number of commercial electronic spectrum analyzers, including the one used by the author to take much of the data presented in this book.

---

<sup>6</sup>In Heisenberg’s formulation, the energy of the electron (or other particle) is given by  $\mathcal{E} = h\nu$  where  $\nu$  is a frequency corresponding to the de Broglie wavelength and  $h$  is Planck’s constant. Hence, in Heisenberg’s formulation of the Uncertainty Principle,  $\Delta\mathcal{E}\Delta t \approx h/2\pi$ .

<sup>7</sup>My personal suspicion is that the peculiar nomenclature arose as a typographical error, compounded by the fact that there actually was an electrical engineer at the Bell Laboratories named Richard W. Hamming who developed his own window function for numerical analysis that is called the ‘‘Hamming Window’’ in the literature.

## 2.14 The Fast Fourier Transform

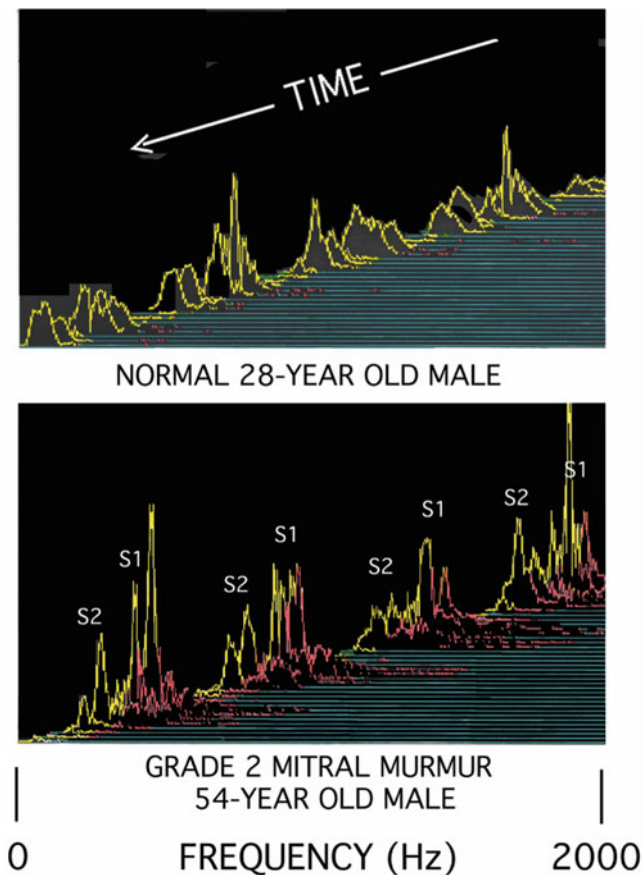
The main problem in applying the straightforward Discrete Fourier Transform to the analysis of data is that the running time for the calculation increases as  $n^2$  where  $n$  is the number of data points to be analyzed. Because one often wants to analyze waveforms consisting of 1000 points or more (for example, one convenient block size is  $2^{10} = 1024$  points), running time is of major importance. Methods to reduce the running time by making use of the redundancy contained in the sine function date at least to the early work of Runge (1903). Most current processors use something known as the fast Fourier transform (FFT) algorithm devised by Cooley and Tukey in 1965. The Cooley–Tukey algorithm reduces the running time from an  $n^2$ -dependence on the number of data points to one that goes up as  $n \log 2n$ . For  $n = 2^{10}$ , that saving can reduce the running time for computer analysis by a factor of 100. (See Brigham and Murrow 1967.) With the advent of high-speed, hardwired FFT processors, it is now possible to do spectral analysis over the entire audio band in real time.

Within the limits imposed by Eq. (2.15), one can use the FFT to study spectra as a function of time. That not only has broad applicability to the study of musical instrument sound generation, but to numerous other areas of science—especially, to medical diagnostics. For example, the FFT is an essential tool for unfolding the data in magnetic resonance imaging (or MRI.) It also can be applied to acoustic diagnostics in medicine. As an example, the variation of the acoustic spectrum of heart sounds with time can be used to diagnose and categorize heart murmurs. (See Bennett and Bennett 1990.)

Figure 2.17 illustrates this technique using the sound monitored by a high-quality condenser microphone placed on the chest at the apex of the heart. The top figure is for a normal 28-year-old male where the spectrum is concentrated below 200 Hz and shown in yellow. The lower figure is for a 54-year-old patient with prolapse of the mitral valve.<sup>8</sup> The data are presented here as a three-dimensional surface in which frequency runs horizontally from near DC to 1000 Hz (left to right) and time advances diagonally from the upper right to the lower left in increments of 0.1 s. The amplitudes of the Fourier components are plotted vertically. Before analysis, the signal was run through an “A-Weighting” Filter that fell off at the low-frequency end so as to mimic the response of the human ear. Hence, what one sees in the figure corresponds to what one would hear through a stethoscope, except that the electronic technique is far more sensitive. Four heart beats are shown in the figure. The signal running from about 200 to 1000 Hz in between the “first” ( $S_1$ ) and “second” ( $S_2$ ) heart sounds from each beat and shown in red is due to the murmur. (The signal below 200 Hz was fairly normal.)

---

<sup>8</sup>Mitral prolapse (verified in the present case by an echocardiogram and open-heart surgery) is a common condition in which the mitral valve (so-called because it is shaped like a Bishop’s mitre) is pushed backward toward the left atrium when the left ventricle contracts. Some blood from the ventricle is then forced back through the leaky mitral valve into the atrium in turbulent flow, instead of going out through the aortic valve in laminar flow, as in the normal case.

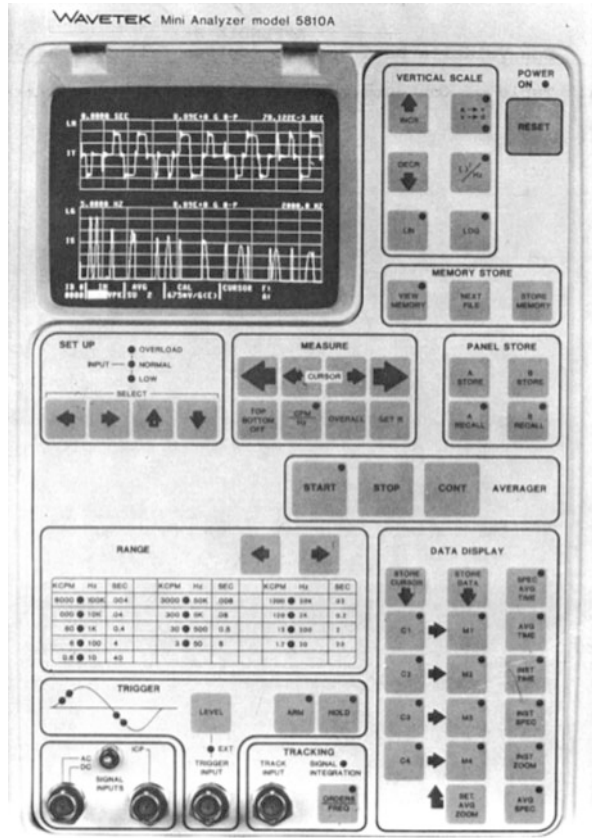


**Fig. 2.17** Spectral surfaces of stethoscopic heart sounds. Upper figure: normal heart sounds from a 28-year-old male. Lower figure: heart murmur arising from mid-systolic mitral prolapse in a 54-year-old male. The RMS (root-mean-square) amplitude is shown vertically and the frequency scale runs from near 0 to 2000 Hz. Time advances diagonally in the plot in increments of 0.1 s. Source: Bennett (1990). The author is indebted to Dr. Lawrence Cohen for helpful discussions

As can be seen from Fig. 2.17, the murmur peaks in the middle of systole (during contraction of the heart) and has its strongest components at about 600 Hz. The murmur was generated by turbulent flow of blood back through the mitral valve into the left atrium when the heart contracted. In contrast to most musical instruments, the relative phases of the spectral components are random. The sound of the murmur is actually very similar to that produced by an African percussion instrument called “The Lion’s Roar.” (It is also akin to the noise made by a crosscut hand saw going through a piece of wood.)

One shortcoming of single FFT-based analysis is that the use of a window function precludes the possibility of reconstructing the original waveform exactly, since part of the information contained in the original waveform is discarded. That

**Fig. 2.18** A real-time FFT (fast Fourier transform) analyzer



is not a problem when one merely wants to determine the main spectral features. However, in cases where one might want to manipulate the data in the frequency domain and then reconstruct a signal in the time domain, that limitation can be a problem. Although one can get around that difficulty by using two FFT processors having identical time windows staggered by half their common duration, the method is cumbersome. A much-touted recent development called Wavelet Analysis appears to offer a more mathematically elegant solution. There, one devises a complete set of wavelet functions that look somewhat like windowed Fourier transform integrands. Since a complete set is involved, the original signal can be reconstructed. (See Rioul and Vetterli 1991.) However, there is a disadvantage in the wavelet analysis method for our present purposes in that the measured frequency intervals increase in powers of two. Although that tends to mimic the logarithmic frequency response of the human ear, such a logarithmic display makes it much harder to pick out the fundamental frequency visually from the spectra of periodic waveforms. In a linear display based on FFT analysis, the harmonic terms are separated by a constant which is usually equal to the fundamental frequency. Much of the data presented in the present book were taken with a real-time, hard-wired FFT analyzer using a “Hanning Window.” (See Fig. 2.18.)



## Problems

**2.1** Suppose the amplitude of the rod vibration in Michelson's spectrum analyzer decayed to  $1/e$  of its initial value in about 5 s. What would the minimum frequency width be that the analyzer could resolve?

**2.2** (a) If the narrowband filter in Fig. 2.3 were 50 Hz wide, what would be the least time you would need to scan through the spectrum from 0 to 10 kHz without distorting the data? (b) What would the frequency scanning rate be in Hz/sec?

**2.3** Draw the amplitude spectrum for the first 15 harmonics of a sawtooth and of a squarewave.

**2.4** If the paper used on an early spectrum analyzer can only be darkened in intensity by a factor of ten, what is its maximum dynamic range?

**2.5** Citizens of Leyden, Massachusetts reported hearing the cannons at Bunker Hill some 80 miles away. Suppose that the sound level in Leyden was about 60 dB ("normal conversation at 3 ft" from Table 2.1). What would the sound level have been 10 ft from a cannon? (Assume the sound pressure falls off as the inverse square law.) [Reference: Arms 1959.]

**2.6** When the first atomic bomb was exploded at the "Trinity" test site in New Mexico, Robert Serber (Oppenheimer's assistant at Los Alamos) heard the sound of the explosion about a minute and a quarter after the flash. How far was he from the explosion? Noting that the blast was heard at Los Alamos 20 min after the flash, how far was Los Alamos from the test site? Use the value for the velocity of sound at  $0^\circ$  in dry air from Table 1.2. About how many dB louder would the blast have been at Serber's location than at Los Alamos? (Data from Serber and Crease 1998, pp. 91, 93).

**2.7** The loudest natural noise in recorded history is said to have occurred on August 27, 1883 when the volcanic island Krakatoa blew up. The sound was heard on Rodriguez Island 3000 miles away. If the level there were 60 dB ("normal conversation" at 3 ft), what would it have been one mile away on the island of Verlaten? (Assume the inverse square law.<sup>9</sup>) [Reference: Winchester 2003.]

**2.8** An A-to-D converter samples a microphone voltage with 10-bit accuracy. What is its limiting dynamic range in dB? (Note:  $2^{10} = 1024$ .)

**2.9** A CD recording uses 16-bit samples. About how big a dynamic range in dB would it provide?

**2.10** A large gymnasium is to be constructed in the middle of a residential area in Bryn Mawr, PA with an air conditioning unit installed on the roof that will produce

---

<sup>9</sup>The assumption of an inverse square law is not terribly good here because the source of noise extended over a distance larger than the island of Verlaten. Also, the radiation pattern appears to have a strong dipole component. (See Winchester 2003, p. 271.)

a sound level of about 80 dB at a distance of 50 ft. Assuming an inverse square law loss, what intensity will this sound have at a neighbor's house 300 ft away?

**2.11** Suppose four identical air conditioning units were to be placed on the roof of the gymnasium in Problem 2.7). (a) What would the increase in sound level be if the four sources were in phase? (b) What would it be on the average if the four phases were randomly related?

**2.12** The air conditioning unit in the previous problem produced the following spectrum, as measured by the instrument in Fig. 2.4: What was the amplitude spectrum?

Frequency (Hz)	31.5	63	125	250	500	1000	2000
Signal (dB)	65	78	88	80	70	67	50

**2.13** The emergence of 17-year cicadas on the weekend of May 22, 2004 resulted in the following spectrum at noon measured in the author's backyard at Haverford, PA: Draw the amplitude spectrum.

Frequency (Hz)	31.5	63	125	250	500	1000	2000	4000	8000
Signal (dB)	50	52.5	52.5	55	60	67.5	60	57.5	52.5

**2.14** We know the harmonics of a square wave decrease in amplitude as  $1/n$  (with  $n$  odd). Draw a spectrum through  $n = 11$  of a square wave in dB referred to the value at  $n = 1$ .

**2.15** Sketch waves proportional to  $\sin \theta$  and  $0.3 \sin 3\theta$  over the range  $0 \leq \theta \leq 2\pi$ . Then, add the two together and sketch the resultant waveform. What might produce that waveform?

**2.16** Suppose the wind supply to an organ pipe were modulated in amplitude sinusoidally at 6 Hz. What would the effect be on the sound spectrum? (Hint: Note the trig identity  $\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B$  and take into account that each harmonic component from the sound wave is multiplied by a sinusoidal term at 6 Hz.)