Additive Synthesis

Additive synthesis is a class of sound synthesis techniques based on the summation of elementary waveforms to create a more complex waveform. Additive synthesis is one of the oldest and most heavily researched synthesis techniques. This section starts with a brief history of additive synthesis and explains its fixed-waveform and time-varying manifestations. The next section is devoted to the process of analysis/resynthesis—linking an analysis of a sound to a resynthesis stage based on additive synthesis.

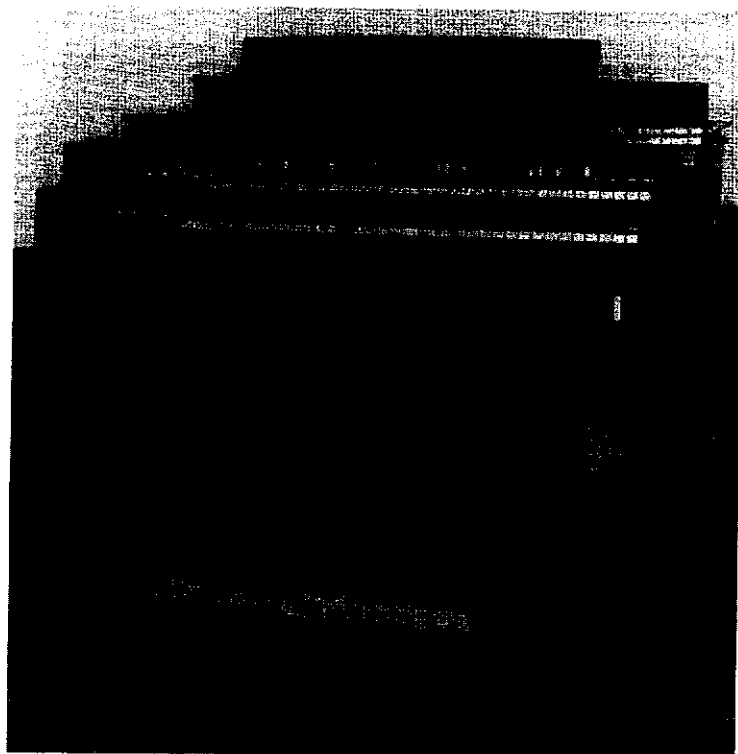### Additive Synthesis: Background

The concept of additive synthesis is centuries old, first being applied in pipe organs by means of their multiple *register-stops*. By pulling on a register-stop, air could be routed to a set of pipes. The air was actually released into the pipe—creating sound—by pressing a key on the organ keyboard. By pulling several register-stops in various proportions one could add together the sound of several pipes for each key pressed on the organ musical keyboard. According to one scholar, "The Middle Ages particularly favored the 'mixtures' in which every note was accompanied by several fifths and octaves based upon it" (Geiringer 1945). This idea of frequency "mixtures" is the essence of additive synthesis.

Additive synthesis has been used since the earliest days of electrical and electronic music (Cahill 1897; Douglas 1968; *die Reihe* 1955; Stockhausen 1964). The massive Telharmonium synthesizer unveiled in 1906 summed the sound of dozens of electrical tone generators to create additive tone complexes (figure 4.12).

Incorporating a miniature version of the Telharmonium's rotating tone generators, the famous Hammond organs were pure additive synthesis instruments (figure 4.13). The power of additive synthesis derives from the fact that it is theoretically possible to closely approximate any complex waveform as a sum of elementary waveforms. Methods exist for analyzing a sound such as a violin tone and resynthesizing it using time-varying combinations of sine waves of various frequencies, phases, and amplitudes. Due to intrinsic limitations in the resolution of the analysis, however, this reconstructed version is never a sample-for-sample replication of the original signal (see chapter 13).

Any method that adds several elementary waveforms to create a new one could be classified as a form of additive synthesis. For example, some forms of granular synthesis discussed in chapter 5 could be called additive synthe-

Figure 4.12 Additive synthesis of a complex tone in the Telharmonium. Sine wave harmonics from the tone-generating alternator are fed to bus bars (54). Pressing a key (C in this case) connects each harmonic to a multicoil transformer (56 "inductorium") where they mix. Each harmonic is attenuated to the desired level by the inductors in series with each winding (56a, b, etc.). The tap-switch inductors (60) regulate the amplitude of the mixing transformer output, as do the inductors near the loudspeakers (72, 73) at the listener's end of the transmission line. (Cahill patent drawing, reproduced in Johnson et al. 1970.)

Figure 4.13   Hammond B3 organ, an additive synthesis instrument based on electromechanical tone-wheels. Different mixtures of the various harmonics can be adjusted by pulling "drawbars" above the musical keys. (Photograph courtesy of the Institute of Organology, Kunitachi College of Music, Tokyo.)

sis techniques (Risset and Wessel 1982). However, we have separated these techniques from additive synthesis in this chapter in order to clarify the distinctions between the classical method of sine wave additive synthesis and those methods.

Fixed-waveform Additive Synthesis

Some software packages and synthesizers let the musician create waveforms by *harmonic addition*. In order to make a waveform with a given spectrum the user adjusts the relative strengths of a set of *harmonics* of a given fundamental. (The term "harmonic" as an integer multiple of a fundamental frequency was first used by Sauveur [1653–1716] in 1701.) For example, 400 Hz is the second harmonic of 200 Hz, since 2 times 200 equals 400. The harmonics can be displayed as a *bar graph* or *histogram*, with the height of each bar representing the relative strength of a given harmonic. Figure 4.14 shows a harmonic spectrum and the corresponding waveform.
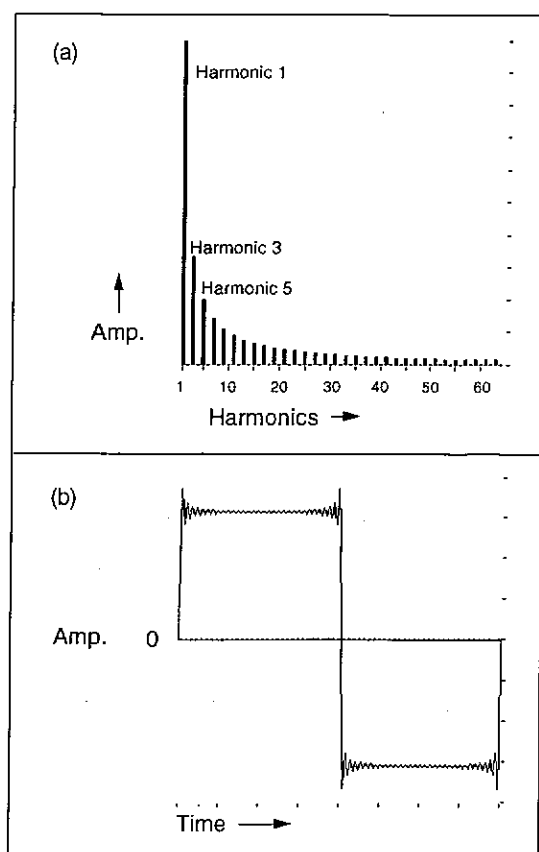
Figure 4.14 Waveform synthesis by harmonic addition. (*a*) Histogram showing the relative strength of the harmonics on a linear scale. In this case, the histogram has energy only in the odd harmonics. The amplitude of the third harmonic is one-third that of the fundamental, the amplitude of the fifth harmonic is one-fifth that of the fundamental, and so on. (*b*) Approximation to a square wave synthesized by harmonic addition using the histogram in (*a*).

Once a desirable spectrum is tuned, the software calculates a waveform that reproduces the spectrum when it is played by a digital oscillator. This spectrum template aligns to different frequencies when one changes the pitch of the oscillator. Figure 4.15 shows successive stages of waveform addition used to create a quasi-square wave.

The Phase Factor

Phase is a trickster. Depending on the context, it may or may not be a significant factor in additive synthesis. For example, if one changes the starting phases of the frequency components of a fixed waveform and resynthesizes the tone, this makes no difference to the listener. And yet such a
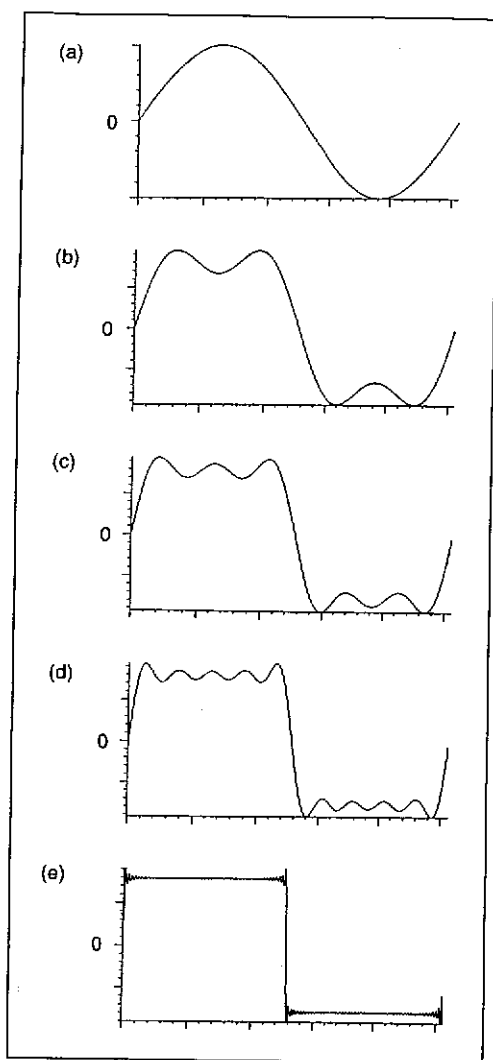
Figure 4.15   Stages of harmonic addition as seen in a series of time-domain waveforms. (*a*) Fundamental only. (*b*) First and third harmonics. (*c*) Sum of odd harmonics through the fifth. (*d*) Sum of odd harmonics through the ninth. (*e*) Quasi-square wave created by summing odd harmonics up to the 101st.
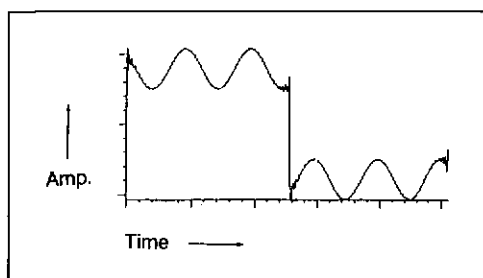
Figure 4.16 Effect of phase in additive synthesis. This waveform is the result of the same mixture of sine waves as in figure 4.15e except that the starting phase of the fifth harmonic is 90 degrees instead of 0 degrees.

change may have a significant effect on the visual appearance of the waveform, as shown in figure 4.16.

Phase relationships become apparent in the perception of the brilliant but short life of attacks, grains, and transients. The ear is also sensitive to phase relationships in complex sounds where the phases of certain components are shifting over time. As we see later in the section on sound analysis and resynthesis, proper phase data help reassemble short-lived components in their correct order, and are therefore essential in reconstructing an analyzed sound.

Addition of Partials

We can generalize from addition of harmonics to addition of *partials*. In acoustics, a *partial* refers to an arbitrary frequency component in a spectrum (Benade 1990). The partial may or may not be a harmonic (integer multiple) of a fundamental frequency f. Figure 4.17a shows a spectrum containing four partials: two harmonic and two *inharmonic*. An inharmonic partial is not in an integer ratio to the fundamental frequency. Figure 4.17b is the waveform that results from the sum of the four partials.

Addition of partials is limited in that it succeeds only in creating a more interesting fixed-waveform sound. Since the spectrum in fixed-waveform synthesis is constant over the course of a note, partial addition can never reproduce accurately the sound of an acoustic instrument. It approximates only the *steady-state* portion of an instrumental tone. Research has shown that the *attack* portion of a tone, where the frequency mixture is changing on a millisecond-by-millisecond timescale, is by more useful for identifying traditional instrument tones than the steady-state portion. In any case, a time-varying timbre is usually more tantalizing to the ear than a constant spectrum (Grey 1975).
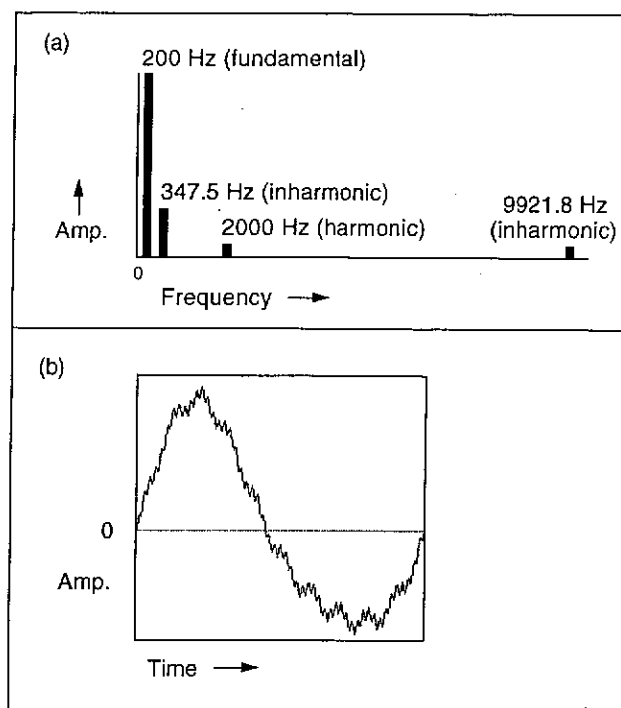
Figure 4.17 Partial addition with four components. The percentage contribution of each component is 73, 18, 5, and 4 percent, respectively. (*a*) Frequency-domain view. (*b*) Time-domain waveform.

Time-varying Additive Synthesis

By changing the mixture of sine waves over time, one obtains more interesting synthetic timbres and more realistic instrumental tones. In the trumpet note in figure 4.18, it takes twelve sine waves to reproduce the initial attack portion of the event. After 300 ms, only three or four sine waves are needed.

We can view the process of partial addition graphically in several ways. Figure 4.19a shows additive synthesis in the analog domain, as it was practiced in the 1950s (Stockhausen 1964). The figure shows several oscillator hardware modules, each with a manually controlled frequency knob. The outputs of the oscillators are routed to a mixer. The composer adjusted the balance of the oscillators in real time to determine the time-varying spectrum. With this setup, manual control was the only option. To precisely realize a time-varying mixture took several people working together (Morawska-Büngler 1988).

Figure 4.19b shows digital additive synthesis. An audio oscillator is represented as a quasi half-circle with a pair of inputs—one for amplitude and
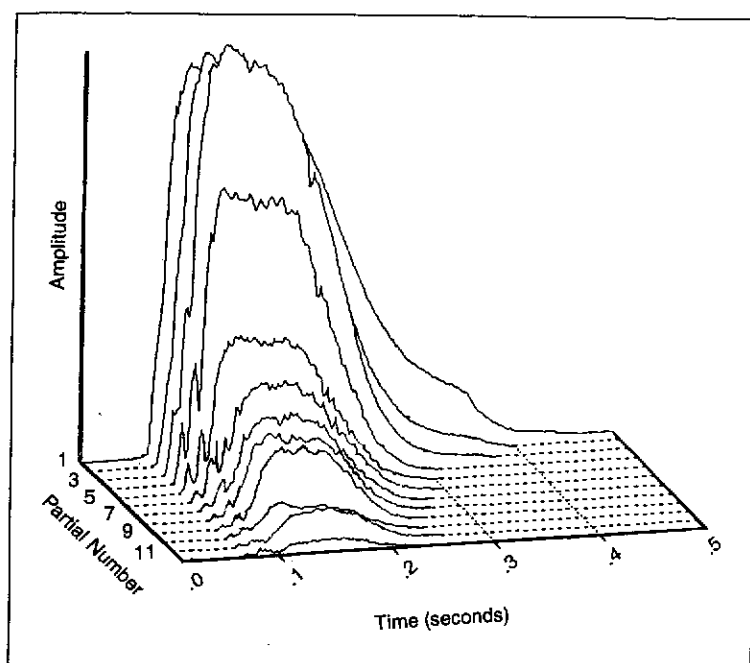
Figure 4.18 Time-varying spectrum plot of twelve partials of a trumpet tone, with the highest partials in the foreground. Time goes from left to right. Notice that the fundamental (at the back) is not the highest amplitude, but it lasts the longest.

one for frequency. To generate a time-varying spectrum, each frequency and amplitude input to the oscillators is not a constant but a time-varying envelope function read over the duration of the event. The sine wave audio oscillators feed into a module that sums the signals. The sum module then passes the additive result to a DAC for conversion to sound.

Demands of Additive Synthesis

Time-varying additive synthesis makes heavy demands on a digital music system. First, it requires large numbers of oscillators. If we make the musically reasonable assumptions that each sound event in a piece may have up to 24 partials (each generated by a separate sine wave oscillator), and that up to sixteen events can be playing simultaneously, we need up to 384 oscillators at any given time. If the system is running at a sampling rate of 48 KHz, it must be capable of generating 48,000 × 384 or 18,432,000 samples/second. Since each sample requires about 768 operations (multiply-adds), the total computational load is over 1.4 billion operations per second, without counting table-lookup operations. Table 20.1 in chapter 20
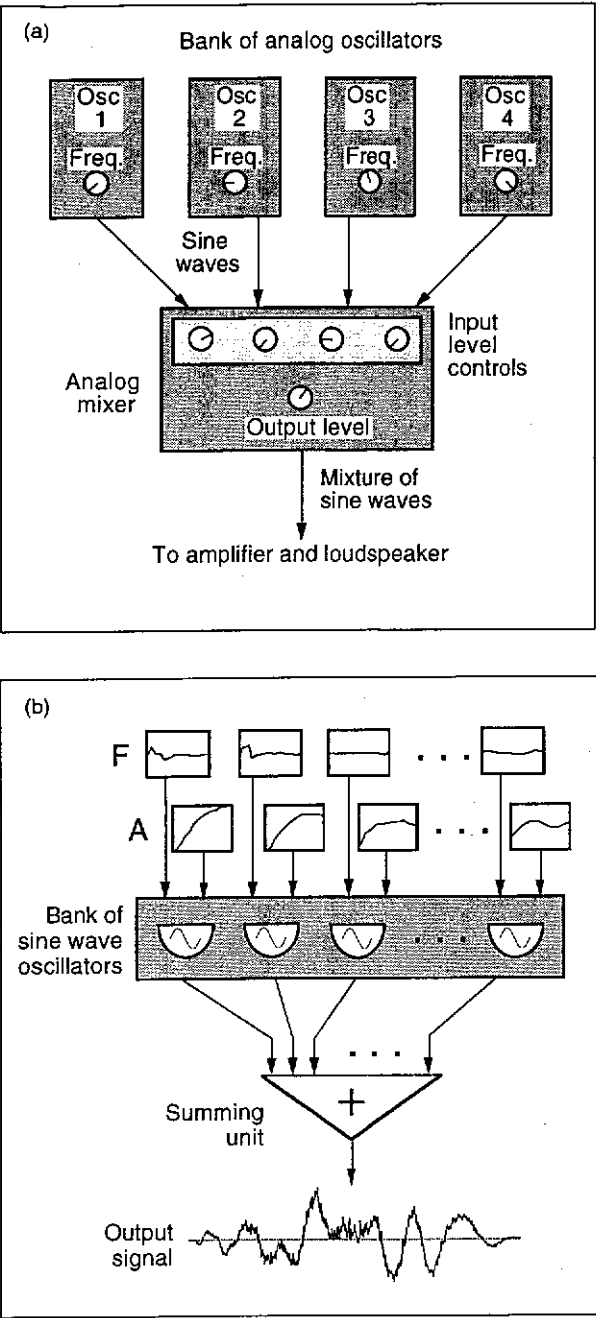
Figure 4.19   Two views of additive synthesis. (*a*) In the analog domain, oscillators feed a mixer. (*b*) Digital additive synthesis. Time-varying additive synthesis with separate frequency (F) and amplitude (A) envelopes. Figure 3.10 shows a more detailed instrument diagram for additive synthesis.

estimates the per-sample requirements for additive synthesis. Such computational demands, although formidable, are not outside the limits of current hardware. For example, one synthesizer specialized for additive synthesis offers the potential of several thousand sine waves in real time (Jansen 1991).

Yet computational power is only one requirement of additive synthesis. The method also has a voracious appetite for control data. If a piece contains 10,000 events (such as a typical orchestral score), each with up to 24 partials, one needs to have 240,000 frequency envelopes and 240,000 amplitude envelopes on hand. Even if the same envelope is used in more than one event, where do the control data come from? This is the subject of the next section.

Sources of Control Data for Additive Synthesis

Effective use of any digital synthesis technique—including additive synthesis—depends on having good control data for the synthesis instrument. To create animated sounds with a rich internal development, one drives the synthesizer with control data; hence, the control data are also referred to as the *driving functions* of the synthesis instrument. Control data can be obtained from several sources:

1. Imported from another domain and mapped into the range of synthesis parameters. For example, some composers have traced the shape of mountains or urban skylines and used these curves as control functions. This is the approach used in the early computer music piece *Earth's Magnetic Field* (1970) by Charles Dodge and in pieces derived purely from geometric, stochastic, or other mathematical or physical models.

2. Generated by a composition program that embodies composer-specified constraints on musical microstructure. An example is John Chowning's *Stria* (1977), realized with additive synthesis of inharmonic spectra.

3. Generated by an interactive composition system that translates high-level musical concepts such as *phrases* (in the Formes language of Rodet and Cointe 1984), *tendency masks* (as in the POD system of Truax 1977, 1985), *sound objects* (as in the SSSP system of Buxton et al. 1978), or *clouds* (as in asynchronous granular synthesis of Roads 1978c, 1991) into synthesis parameters.

4. Entered manually by the composer, using combinations of the previously mentioned sources or the composer's intuitive, theoretical, or empirical knowledge of psychoacoustics. An example of this method is Jean-Claude Risset's piece *Inharmonique* (1970).

5. Supplied from an analysis subsystem that mulches a natural sound and spews out the control data needed to resynthesize it. The data can also be edited in order to create transformations of the original sounds. Trevor Wishart (1988) used sound analysis as an intermediate stage in transforming vocal sounds for his piece *Vox-5* (see also Murail 1991).

Since methods 1 to 4 are based on compositional aesthetics, we do not discuss them further in this chapter. The fifth method requires a subsystem for sound analysis; this is the subject of the next section.
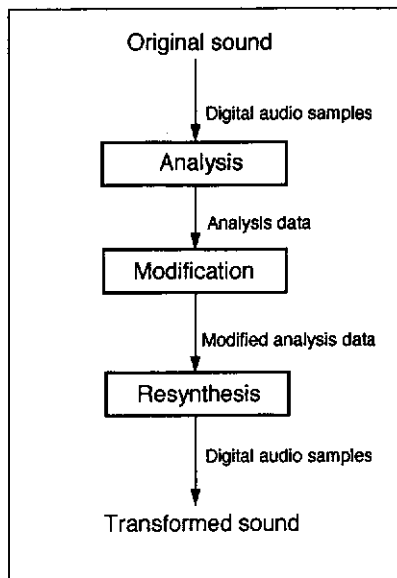
Additive Analysis/Resynthesis

Analysis/resynthesis encompasses different techniques that have a three-step process in common (figure 4.20):

1. A recorded sound is analyzed

2. The musician modifies the analysis data

3. The modified data are used in resynthesizing the altered sound

The concept of analysis/resynthesis is not predicated solely on additive synthesis. It can also be based on subtractive resynthesis (see chapter 5),

Figure 4.20   General overview of analysis/resynthesis. The modification stage may involve manual edits to the analysis data or modifications via *cross-synthesis* where the analysis data of one sound scale the analysis data from another sound.

combinations of additive and subtractive resynthesis (Serra 1989; Serra and Smith 1990), or other methods (see chapter 13).

Early experiments in additive analysis/resynthesis were carried out by H. Fletcher (of the famous Fletcher-Munson loudness curves) and his associates (Fletcher, Blackham, and Stratton 1962; Fletcher, Blackham, and Christensen 1963). They used entirely analog equipment. When digital additive methods are used for resynthesis, the entire system looks like figure 4.21. The analysis is carried out successively on short segments of the input signal. The process of segmenting the input signal is called *windowing* (discussed in chapter 13 and appendix A). We can think of each windowed segment as being sent through a bank of narrow bandpass filters, where every filter is tuned to a specific center frequency. In practice, a *fast Fourier*
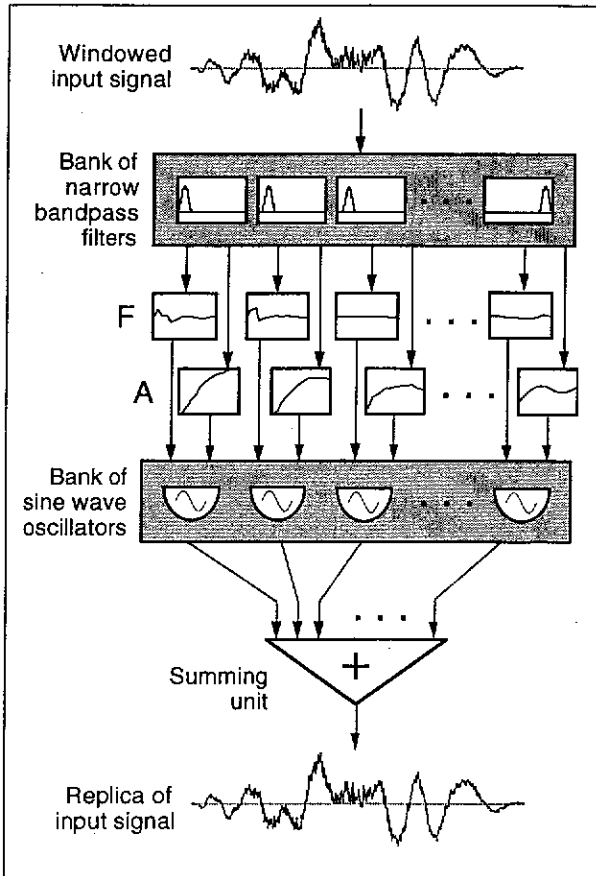
Figure 4.21 Additive analysis/synthesis. A windowed input signal is analyzed by a filter bank into a set of frequency (F) and amplitude (A) envelopes or *control functions* that drive a set of oscillators. If the analysis data are not changed, the output signal should be almost the same as the input signal.

transform (FFT) usually replaces the filter bank and performs essentially the same task in this application, that is, measuring the energy in each frequency band (again, see chapter 13 and appendix A).

The amplitude of the signal coming out of each filter is measured, and this time-varying value becomes the *amplitude control function* for that frequency range. At the same time, the system calculates control functions corresponding to small frequency deviations by looking at the output of adjacent filters (or *analysis bins,* in the case of the FFT).

The frequency and amplitude control functions drive a bank of oscillators in the resynthesis stage. In other words, we are using the information gleaned from the analysis of an existing sound to create the set of control functions needed to resynthesize that sound additively with sine waves. If the input sound is well modeled as a sum of sine waves, the summed signal generated by the oscillators should be much the same as the original input signal.

Of course, straightforward analysis/resynthesis of a sound is not interesting from a musical standpoint. In order to create musically interesting effects, we must modify the data generated by the analysis. This is the subject of the next section.

Musical Applications of Additive Analysis/Resynthesis

After the analysis has been performed, musicians can edit the control functions to create variations of the original input signal. Many different effects are possible with this technique, as listed in table 4.1. Three compositions produced in the 1980s stand as good examples of compositional manipulation of analysis data: *Mortuos Plango, Vivos Voco* (1981) by Jonathan Harvey, *Désintegrations* (1983, Salabert Trajectoires) by Tristan Murail, and *Digital Moonscapes* (1985, CBS/Sony) by Wendy Carlos.

In the Harvey piece, the composer analyzed the sound of a large bell. For each sinusoidal component in the resynthesis, the composer substituted the sound of a sampled boy's voice at the appropriate frequency. The voice samples followed the analyzed frequency and amplitude control functions of the chiming bells, creating an eerie effect of a boy/bell chorus. In the composition by Murail, the composer analyzed traditional instrument tones and created synthetic complements to these tones that blend seamlessly with the sounding instruments yet weave out dramatically when the instruments stop. *Désintegrations* is a classic example of *spectral composition* techniques where the harmonic structure of the work is based on an analysis of instrumental tones (Murail 1991). In *Digital Moonscapes,* Carlos used analysis data as the inspiration for creating an ad hoc synthetic orchestra of per-

Table 4.1 Musical transformations using additive analysis/resynthesis

Musical effect	Technique
Variations of recorded sounds	Change selected frequency or amplitude envelopes by editing or multiplications by arbitrary functions.
Spectrum scaling (without time scaling)	Multiply the frequency of all the partials (possibly excepting the fundamental) by a factor n or by arbitrary functions. Since multiplication does not preserve formant structures, vocal and instrumental sounds may lose their characteristic identity.
Spectrum shifting (without time scaling)	Add a factor n or an arbitrary function to all partials (possibly excepting the fundamental). For small values this preserves formant structures.
Spectrum inversion	Reversing the order of the frequency components before resynthesis, so that the amplitude of the first partial is assigned to the last partial, and vice versa, followed by exchange of the amplitudes of the second and next-to-last components, etc.
Hybrid timbres	Replace some envelopes from one sound with selected envelopes from another sound.
Time expansion and compression without pitch shifting	Extend the duration of the frequency and amplitude envelopes, or change the *hop size* on playback (see chapter 13).
Stretch a percussive timbre into a prolonged synthetic passage	Delay the onset time of each partial and smooth their envelopes.
Timbral interpolation from one instrumental tone to another	Interpolate over time between the envelopes of two instrument tones.
Mutating synthetic sounds	Interpolate between the envelopes of arbitrary synthetic sounds.
Enhance the resonance regions of recorded sounds	Increase the amplitude of selected frequency partials.
Cross-synthesis	Method 1: Use the amplitude envelopes for the partials of one sound to scale the amplitude envelopes of another sound (see *fast convolution* in chapter 10).
	Method 2: Apply the amplitude envelopes from one sound to the frequency (or phase) functions of another sound.
	Method 3: Apply the noise residual from one sound to the quasi-harmonic part of another sound (see, for example, the description of spectral modeling synthesis and the comb wavelet transform in chapter 13).

cussion-, string-, woodwind-, and brasslike timbres, used in an idiomatic orchestral style.

The next section briefly discusses current techniques of sound analysis with additive resynthesis with an emphasis on the data reduction problem. It serves as a prelude to the more detailed treatment in chapter 13 and appendix A.

Methods of Sound Analysis for Additive Synthesis

Many spectrum analysis methods, including *pitch-synchronous* analysis (Risset and Mathews 1969), the *phase vocoder* (Dolson 1983, 1986, 1989b), and *constant-Q* analysis (Petersen 1980; Schwede 1983; Stautner 1983), among others, are variations on the basic technique of Fourier analysis of component frequencies. The practical form of Fourier analysis is the *short-time Fourier transform* (STFT). This method can be thought of as analyzing a sampled sound by extracting successive short-duration overlapping segments (shaped by a window function) and applying a bank of filters to the selected segment. The output of each filter is measured, indicating the amplitude and the phase of the spectrum at that particular frequency. A series of these short-time analyses (akin to the frames of a film) constitute a time-varying spectrum. At the core of the STFT is the FFT, a computationally efficient implementation of Fourier analysis (Cooley and Tukey 1965; Singleton 1967; Moore 1978a, 1978b; Rabiner and Gold 1975).

The phase vocoder (PV) (Flanagan and Golden 1966; Portnoff 1978; Holtzman 1980; Moorer 1978; Dolson 1983; Gordon and Strawn 1985; Strawn 1985b) deserves special mention here, as it is a popular method of sound analysis/resynthesis that has been distributed with several music software packages. The PV converts a sampled input signal into a time-varying spectral format. In particular, it generates a set of time-varying frequency and amplitude curves. Many interesting sound transformations can be achieved by editing and resynthesizing PV data. For example, the phase vocoder can be used for *time compression* or *time expansion* without pitch change. In this effect a sound is made longer or shorter without significantly affecting its pitch or timbre. (See chapter 10 for a discussion of various approaches to time compression/expansion.)

Contrary to the expectations of the researchers who invented them (who were searching for efficient coding techniques), sound analysis techniques may generate an "information explosion" (Risset and Wessel 1982). That is, the analysis data (the control functions) can take up many times more memory space than the original input signal. The amount of data depends partly on the complexity of the input sound, that is, how many sine wave

functions are needed to resynthesize it, and partly on the internal data representation used in the analysis program. Using the tracking phase vocoder, for example, a short sound file that takes up 2 Mbytes may generate tens of Mbytes of analysis data. Such storage requirements make it difficult to build libraries of analyzed sounds, and the volume of data becomes onerous to edit. This situation mandates some form of data reduction of the control data, the subject of the next section.

Data Reduction in Analysis/Resynthesis

Data reduction is important to efficient analysis/resynthesis. Data reduction takes two steps. First, the data—a set of amplitude and frequency control functions—are analyzed. Second, an algorithm transforms the original data into a more compact representation. An important goal of data reduction is to compact data without eliminating perceptually salient features of the input signal. Another important goal in computer music work is that the analysis data must be left in a form that can be edited by a composer. The goal is not simply to save bits; rather, one wants to make it easy to manipulate the data-reduced material (Moorer 1977).

A large body of research work on data reduction of digital audio samples is recorded in the literature, including studies by Risset (1966), Freedman (1967), Beauchamp (1969, 1975), Grey (1975), Grey and Gordon (1978), Charbonneau (1981), Strawn (1980, 1985a, 1985b), Stautner (1983), Kleczkowski (1989), Serra (1989), Serra and Smith (1990), Holloway and Haken (1992), and Horner, Beauchamp, and Haken (1993). Since real-time work is so important to musicians, one goal of analysis/resynthesis research is to speed up data reduction processing and facilitate real-time synthesis from reduced data. Papers by Sasaki and Smith (1980) and Schindler (1984) explain hardware designs for high-speed digital synthesis from reduced data.

Many volumes of engineering literature explore data reduction methods. Here we glance at four techniques that have been applied in computer music: line-segment approximation, principal components analysis, spectral interpolation synthesis, and spectral modeling synthesis. (See also Goldberg 1989 for a description of the genetic algorithm approach, which has recently been applied to synthesis data reduction [Horner, Beauchamp, and Haken 1993].)

Line-segment Approximation

Line-segment approximation of the amplitude and frequency control functions eliminates the need to store a distinct value for every sample analyzed.

Instead, the analysis system stores only a set of *breakpoint pairs,* which are time (*x*-axis) and amplitude (*y*-axis) points where the waveform changes significantly. Line-segment approximation represents the overall outline of a waveform by storing only the points of maximum inflection (change). In the resynthesis stage the system "connects the dots," usually by means of straight lines interpolated between the breakpoint pairs.

Initial work with line-segment approximation was done by hand, using an interactive graphics editor to construct functions with four to eight segments each (Grey 1975). A data reduction of a hundredfold was achieved. This manual editing work can also be partially automated, as demonstrated by Strawn (1985a, 1985b). Figure 4.22a shows a perspective plot of the sixteen harmonics of a violin tone, sampled at 25 KHz. Figure 4.22b plots an approximation to (a) using just three line segments.

Going beyond the storage of line-segment approximations, Beauchamp (1975) developed a heuristic technique for inferring the approximate amplitude curve of all harmonics of a tone from the curve of the first harmonic. For simple periodic tones, Charbonneau (1981) found that even more radical data reduction could be achieved. He used simple variations of a single envelope for all amplitude functions of a given tone. (See also Kleczkowski 1989 and Eaglestone and Oates 1990 for refinements of these proposals.)

Principal Components Analysis

The technique of *principle components analysis* (PCA) has been applied in several analysis/resynthesis systems (Stautner 1983; Sandell and Martens 1992; Horner, Beauchamp, and Hakken 1993). PCA breaks down a waveform using the mathematical technique of *covariance matrix* calculation. This results in a set of basic waveforms (the principal components) and a set of weighting coefficients for these basic waveforms. When the components are summed according to their weights, the result is a close approximation of the original waveform.

The advantage of PCA is its potential for data reduction. PCA analysis summarizes the underlying relationships between samples so the fewest number of components account for the maximum possible variance in the signal. The process of determining the principal components and their weighting coefficients is implemented as an iterative approximation that tries to minimize the squared numerical error (difference between the original and the approximation). The first principal component is a fit of a single waveform to the entire data set. The second principal component is a fit to the *residual* (sometimes called *residue*), or difference between the original and first approximation. The third component is a fit to the residual of the
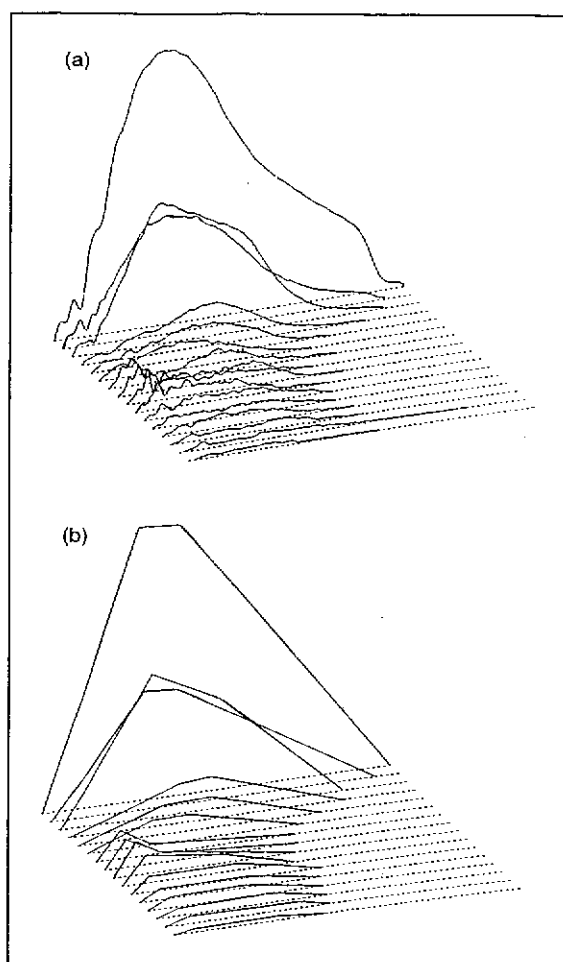
Figure 4.22 Drastic data reduction of analysis data for additive synthesis. Amplitude is plotted vertically, frequency goes from back to front, and time goes left to right. (*a*) Original frequency-vrs-time-vrs-amplitude curve of a violin tone. (*b*) The same violin tone as in (*a*), approximated with only three line segments per partial.

second component, and so on. For further details on PCA see Glaser and Ruchkin (1976).

Spectral Interpolation Synthesis

Spectral interpolation synthesis (SIS) (Serra, Rubine, and Dannenberg 1990) is an experimental technique that generates time-varying sounds by interpolating between analyzed spectra. Rather than crossfading between sampled sounds in the time domain (as in multiple wavetable synthesis discussed in chapter 5), SIS starts from analyses of recorded sounds and

uses additive synthesis to crossfade between the analyses of successive spectra in the frequency domain. An automatic data reduction algorithm is necessary to compress the analysis data into a small set of common spectral paths between two successive sounds and a set of ramp functions that describe the transition from one spectrum to the next. The main difficulty with the procedure appears to be its handling of the attack portion of sounds.
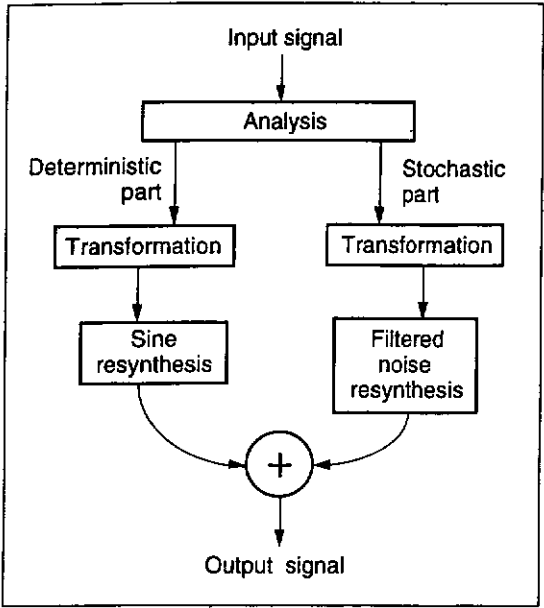
Spectral Modeling Synthesis

Spectral modeling synthesis (SMS) (Serra 1989; Serra and Smith 1990) reduces the analysis data into a *deterministic* component (narrowband components of the original sound) and a *stochastic* component. The deterministic component is a data-reduced version of the analysis that models the most prominent frequencies in the spectrum. These frequencies are isolated by a process of *peak detection* in each frame of the analysis, and *peak continuation,* which tracks each peak across successive frames. SMS resynthesizes these tracked frequencies with sine waves. This is the same method as used in tracking phase vocoders described in chapter 13.

SMS goes beyond this representation, however, by also analyzing the residual or the difference between the deterministic component and the original signal. This is the called the "stochastic" component of the signal. The stochastic component takes the form of a series of envelopes that control a bank of frequency-shaping filters through which white noise is being passed. Thus a composer can transform the deterministic (sine) envelopes and the stochastic (filtered noise) components separately, if desired (figure 4.23). Noisy components remain noisy, even after transformations (such as filtering) are applied to them. This stands in contrast with a pure sine wave model, in which transformations (such as time compression/expansion) on noisy components often turn them into orderly sine wave clusters, denaturing their noisy texture.

Efficient algorithms for generation of *pseudorandom noise* are well known (Knuth 1973a; Keele 1973; Rabiner and Gold 1975). Thus the use of filtered noise results in a tremendous data reduction. In purely sinusoidal resynthesis without this kind of data reduction, noisy components must be approximated with hundreds of sine waves. The control functions for these sine waves take up a great deal of storage space, and the sine wave resynthesis is costly from a computational standpoint.

A problem of accuracy left open by SMS is that the filtered pseudorandom noise it uses to reconstruct the stochastic component is not necessarily the same quality of noise as the original source. In many sounds, "noise" is the result of complicated turbulences that have an audible char-

Figure 4.23 Overview of spectrum modeling synthesis. The input signal is divided into a deterministic part and a stochastic part. Each part can be modified separately before resynthesis. (See figure 13.16 for a more detailed view of the analysis stage.)

acter and identity. For some sounds, the approximation by uniform noise leaves room for improvement.

Walsh Function Synthesis

So far we have discussed analysis/resynthesis as a process based mainly on Fourier analysis with resynthesis based on sine wave summation. The Fourier sine wave approach has a long tradition of research and application stemming from the original theorem that states that for periodic signals, a combination of sine waves of various frequencies can be created that approximate arbitrarily closely the original signal. Mathematical research has shown that other groups of waveforms besides sine waves can be used to approximate signals. A family of square waves called the *Walsh functions* can be used to approximate a signal after it has been analyzed by means of the *Walsh-Hadamard transform*. Walsh functions, being rectangular waves, are a kind of "digital domain series," since they take on only two values $+1$ and -1 (Walsh 1923).

Figure 4.24 presents the first eight Walsh functions. Just as with the Fourier series and its sine waves, an arbitrary periodic waveform can be approximated as an additive sum of a finite series of Walsh functions. While
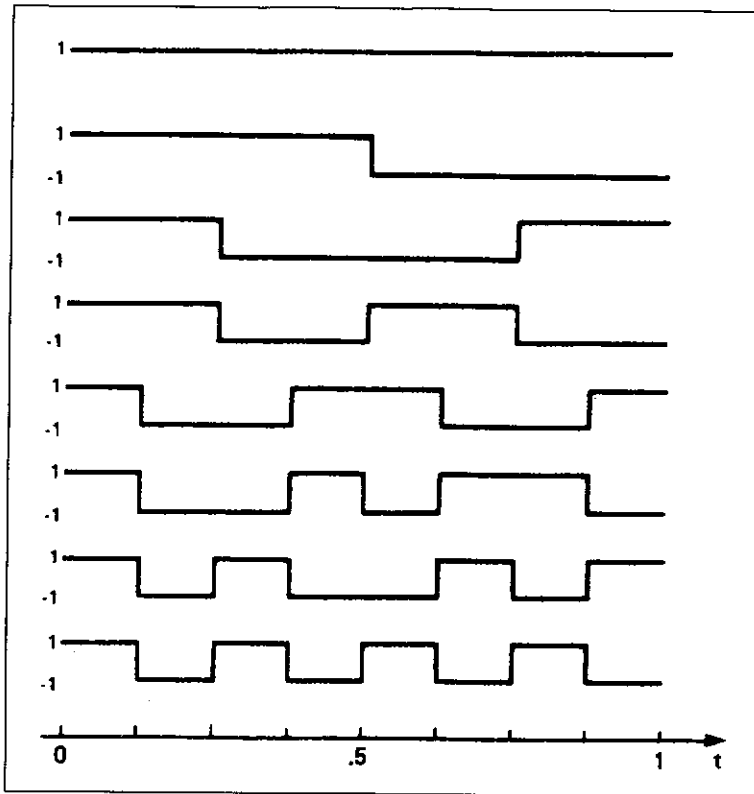
Figure 4.24   The first eight Walsh functions, 0 (top) to 7 (bottom).

the Fourier series builds up waveforms out of component frequencies, Walsh synthesis builds up waveforms using functions of different *sequencies*. Sequency is defined as one-half the average number of zero crossings per second (zps) (Hutchins 1973). Figure 4.25 shows a composite waveform derived by summing several Walsh functions. It suggests how sine wave additive synthesis and Walsh function synthesis are conceptual opposites. That is, the hardest waveform to synthesize in Walsh function synthesis is a pure sine wave. The Walsh approximation to a sine will stay jagged until a very high number of sequency terms are used. Any jaggedness gives an "unsinusoidal" and generally objectionable quality. By contrast, in sine wave synthesis, the hardest waveform to synthesize is one with a rectangular corner, such as a square wave! Figure 4.15, for example, depicts a quasi-square wave constructed by summing 101 sine waves.

The main advantage of Walsh functions in digital sound synthesis is their rectangular shape, a shape that can be computed at high speed by inexpensive digital circuits. A disadvantage of Walsh function synthesis as against
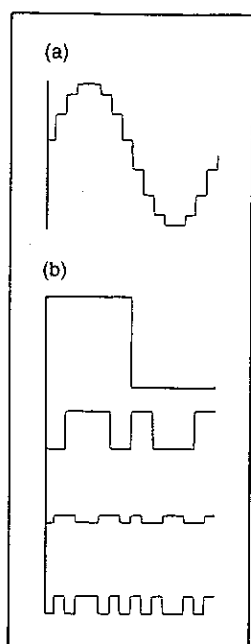
Figure 4.25 Demonstration of Walsh function summation. (*a*) A simple sine wave approximation built by adding the Walsh functions shown in (*b*). (After Tempelaars 1977.)

sine wave synthesis is that individual Walsh functions are not associated with specific harmonics, as they are in sine wave additive synthesis. It is, however, possible to pass mathematically from the Fourier (frequency) domain to the Walsh domain (Tadokoro and Higishi 1978). Thus, one can specify a sound in terms of the addition of various frequency components (partials), and then transform this specification into a set of parameter values for a Walsh function synthesizer. Moreover, natural sounds can be sampled and transformed into the Walsh domain using the Walsh-Hadamard transform and resynthesized using the fast Walsh transform (FWT) (Hutchins 1973, 1975).

A number of music synthesis operations have been redesigned for Walsh signal processing circuits. For example, Hutchins (1973) designed an envelope generator using Walsh function circuits. Rozenberg (1979) and Hutchins (1975) showed how to realize amplitude modulation, subtractive synthesis, frequency modulation, frequency shifting, and reverberation—all in the Walsh domain.

Despite the potential of Walsh function synthesis, only a few experimental devices based on this technique have been built (Hutchins 1973, 1975; Insam 1974). None are commercially available. This is probably due to the

fact that the cost of circuits for sine wave additive synthesis has continued to decline (including memory chips and multipliers), so the economic advantage of Walsh function circuits has diminished. The weight of accumulated research in Fourier/sine wave methods and the more intuitive relationship between frequencies and perception have also contributed to the popularity of sine wave summation in contemporary synthesizer designs.

Conclusion

This chapter has discussed two widely used synthesis techniques: sampling and additive synthesis. The sampler is the mockingbird of musical instruments. Its creative synthesis capabilities may be weak, but it can copy any source through its ability to memorize and playback sound. Because it can also mimic the rich sounds of acoustic instruments, a sampler is among the most popular electronic instruments available.

Additive techniques have been studied in detail for decades. When coupled to an analysis stage, additive synthesis is a powerful means of simulating natural sounds and cloning variations of them. As we have seen, the main drawback of additive techniques is that they achieve a quasi generality by sacrificing computational efficiency. In order to simulate a given sound, a thorough analysis must be performed. This analysis can generate an explosion of data that must go through a data reduction stage before it becomes editable. Sound analysis tools require serious computing power, so in the past they were available only on expensive institutional-grade computing hardware. This situation is changing, with sophisticated sound analysis and editing tools available on even portable computers.

The next chapter discusses two synthesis techniques with links to sampling and additive synthesis, namely multiple wavetable synthesis and granular synthesis. The other technique discussed in chapter 5 is subtractive synthesis, the conceptual opposite of additive synthesis.