

10 Computer Generation and Manipulation of Sounds

STEFANIA SERAFIN

In a 1963 article in the journal *Science*, entitled ‘The Digital Computer as a Musical Instrument’, Max Mathews, the father of computer music, declared the birth of computer-generated sound. He claimed that ‘there are no theoretical limitations to the performance of the computer as a source of musical sounds, in contrast to the performance of ordinary instruments’ (Mathews, 1963).

Sound synthesis can be defined as the production and manipulation of sounds using mathematical algorithms. A useful classification of sound synthesis techniques was given by Julius O. Smith (Smith 1991), who proposed four categories: processed recordings, abstract algorithms, spectral models and physical models. Algorithms in the category of processed recordings require some initial sonic material, rather than generating synthetic sounds from scratch, and include such techniques as wavetable synthesis and granulation. Abstract algorithms directly create sound from general mathematical formulae, such as amplitude, ring, frequency modulation and waveshaping. Spectral models simulate sounds as they are received and perceived by the ear, including techniques such as source-filter synthesis, additive synthesis, the phase vocoder and subtractive synthesis. Smith’s last synthesis category involves physical models, which simulate the source of sound production. We shall consider all of these categories in this chapter.

The Early Days

The first experiments in computer-generated sounds were performed in 1957 at Bell Labs, and remained among a select few into the 1960s. Scientists and musicians certainly did not have our contemporary privilege of being able to synthesise complex sonic patterns in realtime on a personal laptop. At the origins of computer music, only high-end laboratories had the capability to produce sounds by computer, and the generation of a few seconds of sound usually took an hour or more. To complicate the matter, at Bell Labs decks of punch cards with recorded computer music scores had to be carried to the IBM building in Manhattan. A mainframe computer located there then converted the punch cards into a digital sound tape,

[185]

which was then returned to Bell Labs and played back through a digital to analogue converter.

These limitations and challenges certainly did not discourage the pioneers of the field. On the contrary, music software was starting to be developed, among which Music III and its descendants Music IV and V introduced the concept of a unit generator. A unit generator is a building block of a sound synthesis algorithm. Examples of unit generators are oscillators, filters, multipliers and adders, and amplitude envelope generators. Different complex sonic patterns and sound synthesis algorithms could be implemented by connecting different oscillators; see Chapter 4 for more on this background to programming methods for computer music.

Among the computer music pioneers, Jean-Claude Risset, a French composer and scientist, began experimenting with synthetic sounds produced using *additive synthesis*. Additive synthesis is a synthesis technique derived from the Fourier theorem, which states that any periodic function can be formulated as a sum of sine waves. When applied to computer music, the Fourier theorem can be interpreted as the possibility of creating any complex waveform by summing a set of sinusoidal components. A sine wave is produced by a simple oscillator, whose frequency, amplitude and phase can be varied; in additive synthesis, a bank of sine tones, typically at fixed frequencies and starting phases, but with time-varying amplitude envelopes, are summed up to create complex sound, the technique therefore requiring many control parameters.

In 1964, after reading Mathews' paper on the possibility of generating sounds by computers, Risset decided to visit Bell Labs, where he began to investigate the timbre of trumpets using analysis and synthesis techniques in Music IV. Risset discovered some important timbral properties of musical instruments, such as the fact that the attack is essential to recognise the sound of a trumpet. Moreover, by playing a piano sound backwards, he discovered that the spectral description of an instrument is not enough to recognise its timbre. He produced the first synthetic bell sounds using additive synthesis, by understanding the importance of the inharmonic spectra of such instruments and the role of each partial's amplitude envelope. Thanks to these discoveries, Risset pioneered the combination of the disciplines of acoustics, sound synthesis and psychoacoustics, where the mathematical understanding of musical sounds and their reproduction by computers are tightly linked to the way such sounds are perceived by humans.

In 1968, after a few years back in France, Risset returned to Bell Labs and created a catalogue of computer-generated sounds, an important research contribution. In this catalogue, guidelines to synthesise different musical instruments using the notation of the Music V program were

provided; particular focus was placed on bell and woodwind sounds. At the same time, Risset produced compositions using sound synthesis, such as the *Computer Suite from Little Boy*, motivated by the Hiroshima bombing. *Little Boy* explores instrumental simulation by additive synthesis, timbral mixing and auditory illusions impossible to generate using acoustical instruments. As an example, Risset used Shepard tones, which create an illusion of never-ending ascending or descending glissandi. The composition clearly shows Risset's interest on the influence of psychoacoustics on computer music, especially concerning the way the sonic structures affect the perception of the resulting sounds.

Another pioneer of computer music research who was strongly influenced by Max Mathews' 1963 paper was John Chowning. Chowning arrived at Stanford as a young graduate student in 1962. As a composer, he had become interested in electronic music after having attended concerts in Paris, especially Pierre Boulez's *Domaine Musicale* series. When a colleague from Stanford handed him a copy of Mathews' paper, Chowning immediately arranged a trip to Bell Labs, to gain a deeper knowledge of the possibilities offered by sound synthesis. In particular, Chowning was intrigued by the sentence in the paper stating that a computer could give unlimited sonic possibilities, as opposed to traditional musical instruments. Returning to Stanford's newly established artificial intelligence lab, which later became the Center for Computer Research in Music and Acoustics (CCRMA), Chowning started to explore the musical potential of computer-generated sounds.

While playing with combining oscillators, he discovered the most commercially successful sound synthesis technique to date: *frequency modulation*, commonly known as FM synthesis. The main idea behind frequency modulation is that when the frequency of one oscillator is modulated by another, a very complex spectrum appears. FM was particularly interesting at that time when computational cost was a real issue: thanks to FM, interesting sounds were produced using the mere combination of two nested sine waves. Chowning's discovery captured Yamaha's attention. The company bought the FM patent and in 1983 released the DX7, the most successful hardware synthesiser in history. As a composer, Chowning naturally used FM in many of his pieces. For example, in *Turenas*, completed in 1972, FM synthesis is used to generate spectral transformations from harmonic to inharmonic spectra. *Turenas* is a four-channel composition, which uses spatialisation algorithms developed by Chowning himself. FM synthesis was also used by composer Paul Lansky in his first computer music piece, *Mild und Leise*, composed in 1973. It is interesting to notice how Radiohead's *Idioteque*, from the album *Kid A* (2000), samples a snippet of this piece.¹

In parallel to the development of new synthesis techniques, engineers and musicians were building new hardware synthesisers which could allow powerful sound manipulations and processing. In 1975, John Appleton produced the prototype for a self-contained digital synthesiser, in association with the New England Digital Corporation, commercially known as the Synclavier. The Synclavier had a bank of timbre generators, each providing a choice of up to twenty-four sinusoidal frequency components for each voice, depending on the version. Its on-board microcomputer had 128 kBytes of memory mainly used for sequencing.

In October 1977, CCRMA acquired the Systems Concepts Digital Synthesizer, commonly known as the Samson Box, named after its designer Peter Samson. The Samson Box, which resembled a big green refrigerator, provided 256 unit generators and 128 different modifiers such as filters, envelope generators or random number generators. Each modifier could be combined with delay units to produce reverberation effects. Moreover, the box provided four analogue-to-digital converters allowing four channels of sound output. All the synthesis techniques known at the time, such as additive, subtractive and FM synthesis were supported on the Samson Box. The box was a clear success, and much music was produced at the time. For example, in 1980 Gareth Loy composed *Nekyia*, a four-channel composition combining recorded and synthesised sounds. Unfortunately this dedicated machine required special support, and lots of effort was put into software and hardware maintenance. Since developing new synthesis algorithms was far from straightforward in the Samson Box, the box became more of a musical instrument than a research tool.

The events described so far all took place in the United States, predominantly at Bell Labs or Stanford University. In 1970 Pierre Boulez was asked to become director of the *Institut de Recherche et Coordination Acoustique Musicque* (IRCAM) in Paris. At the time of its creation, and for some decades hence, IRCAM represented the main centre of research and development in computer music in Europe. Several composers and researchers – including Jean-Claude Risset, David Wessel and Tod Machover, to cite only a few – were invited to Paris to bring their expertise and contribute to the development of the centre.

The initial success of IRCAM was due to the synergy of people working there and the possibilities offered by using powerful technology developed in house, such as the 4X digital synthesiser designed by Peppino Di Giugno. Most of the synthesis techniques described in this chapter were developed or refined at IRCAM. From the 1980s, the development of personal computers facilitated the creation of further research centres around the United States and Europe. Nowadays computer music research and musical creation are spread across many different locations worldwide and accessible to

anyone with even a basic computer. In the following, we further investigate various available synthesis techniques, together with their applications in musical creations.

Granular Manipulation of Sounds

In musical terms, a sound grain can be defined as a short sonic snippet of about ten to a hundred milliseconds, an elementary particle as opposed to a complex soundscape. By combining different grains over time, and by overlapping several grains at the same instant, interesting sonic effects can be produced. The synthesis technique in which different sound grains are combined is known as *granular synthesis*. One of the pioneers of the use of granular synthesis in computer music is Curtis Roads. Working together with his teacher Iannis Xenakis, he investigated the idea of composing with sound particles, an idea mainly inspired by the theory of the Nobel Prize winner Dennis Gabor, who claimed that all sounds can be considered as being made of elementary sound particles limited in time, frequency and amplitude.

In 1974 Roads wrote a computer program with Music V, implementing sound particle synthesis. A succession of such programs and techniques have been developed for his compositions in subsequent decades (Roads 1978; Roads 2001). His latest recorded collection of works is *Point Line Cloud* (2004); a point represents a grain, a line represents a set of points which create a musical tone, and a cloud is a connection of many grains played simultaneously.

Since its conception, many composers have utilised granular synthesis as a musically powerful technique to create and manipulate complex sonic universes using basic particles. Barry Truax has researched granular synthesis extensively, and in 1986 he produced a realtime implementation using a digital signal processor controlled by a microcomputer, his PODX system, creating *Riverrun*. By 1987, Truax was using the technique of granulation to process sampled sounds as compositional material. In *The Wings of Nike* (1987) short sonic grains are preferred, while in pieces such as *Pacific* (1990) longer sequences of environmental sounds are sculpted. In each of these compositions, the granulated material is time stretched by various amounts and thereby produces a number of perceptual changes that seem to originate from within the sound.

A technique which is strongly related to granular synthesis is the so-called Fonction d'onde formantique (*FOF*), or formant wave function, developed at IRCAM in the early 1980s by Xavier Rodet, Yves Potard and Jean-Baptiste Barrière. FOF is used to synthesise vocal sounds by using

short decaying sinusoidal bursts synchronously spaced over time. The resonant frequency of a FOF corresponds to one of the formants (that is, to one of the main resonances) of the vocal tract. By combining several FOFs together, simulation of vocal sounds can be produced. In 1984, the FOF algorithm was used to synthesise Mozart's famous Queen of the Night aria, using a synthesiser called *Chant*. This demonstration was motivated by the need to create some convincing musical examples which utilised the Chant software, in order to show off its musical possibilities. A considerable amount of time was dedicated to carefully synthesising the aria, with an impressive result for that date. However, it is important to note that the team only tackled that section of the aria where vowels predominate, vowels being notoriously easier to synthesise than consonants.

Since the 1990s granular processes have been extensively used² by many composers for the wide musical possibilities they offer, and they are readily available in many different software platforms. Curtis Roads' book, *Microsound* (Roads 2001) provides an overview of many associated techniques and compositions.³

Sound Modelling

Among the different synthesis techniques, sound modelling techniques have seen great interest from acousticians, engineers and computer scientists, and to a lesser extent been taken up by composers. While acousticians are interested in understanding how different musical instruments produce sound, engineers and computer scientists are interested in developing efficient yet accurate algorithms to simulate such sounds, and musicians and composers are interested in using modelling techniques to extend the sonic possibilities offered by traditional instruments.

Sound modelling techniques are commonly divided into *spectral models* and *physical models*. While spectral models simulate how a sound is perceived by the listener, physical models reproduce the source sound production mechanism. An advantage of spectral models is the availability of analysis techniques which allow the obtaining of control parameters for the models from recordings of real instruments. Such analysis techniques do not exist for physical models, so physical models usually also rely on spectral analysis techniques. Physical simulations require a dedicated model for each instrument or sounding object reproduced, while spectral models have a single representation which can then be adapted to different instruments.

It is essential to stress the distinction between the internal representation of a model (the mathematical model being employed to design it) and

how it is seen from the outside (the external representation). Acousticians, engineers and computer scientists are concerned both about internal and external representation. On the other end, from the perspective of a musician, it is especially important that the external representation is understandable and usable. A model with few accessible control parameters can quickly become musically uninteresting, since very limited variations can be introduced. On the other side, a model with too many parameters or whose parameters are not understandable can easily become unusable and too complex. It is an additional challenge for the scientist, and especially for the interaction designer, to find the right trade-off between complexity and musical appeal.

Spectral Modelling

Spectral modelling techniques are perhaps the most popular approach to sound synthesis. By using the Fourier transform, a sound is decomposed into its elementary sinusoidal components. Spectral modelling techniques are a derivation of previous research on additive synthesis as performed in the early years of computer music, together with phase vocoder techniques developed mainly for speech analysis and synthesis. Spectral modelling techniques allow several important sonic manipulations of original material, since every sound can be analysed, transformed in different ways and resynthesised. Common transformations include pitch shifting, time stretching and spectral morphing, the latter being a combination between spectra from different sounds to create hybrid instruments not existing in the real world. The different possibilities offered by spectral models have been for a long time very attractive to computer music composers.

Spectral modelling techniques facilitated the creation of so-called *spectral music*, music concerned with timbral structures obtained by Fourier-based analysis techniques. Originating in France in the 1970s, the 'spectral school' was nurtured by IRCAM, and included such figures as Gérard Grisey and Tristan Murail. Whilst it could include computer-based spectral manipulation and transformation of sound, the computer was also used as an analysis tool, with compositions then developed as scored settings for performance by specially trained musicians, such as Grisey's *Partiels* (1975), for an ensemble of eighteen musicians and based on an analysis of trombone harmonics. As another example, in 1980, Jonathan Harvey received an invitation from Pierre Boulez to work at IRCAM. During this time, Harvey composed *Mortuos Plango, Vivos Voco*, a tape piece that features extensive use of spectral manipulation techniques. Harvey was particularly interested in the technique of spectral morphing, and merged the

sound of the great tenor bell in Winchester cathedral with the sound of his son singing.

Figure 10.1 shows a block diagram of a simple spectral processing framework. In this example the analysis part is rather straightforward, since the input signal is a sine wave. The transformation applied to the sound is a change in frequency, obtained in the frequency domain after calculating the Fourier transform of the original sound.

Research in psychoacoustics has shown that noisy transient components of a sound (often in the attack portion of a sound) are especially important in identifying a particular musical instrument, or in differentiating sound quality between original and synthetically generated instruments. Since the transient portion of a sound is especially hard to synthesise using a sum of sinusoids, sometimes researchers prefer to use a sampled attack rather than a synthesised one. This obviously preserves the quality, but lacks the flexibility offered by richer sound synthesis. In the late 1980s, a technique called *sines plus noise* was developed by Xavier Serra as part of his PhD dissertation at Stanford University (Serra 1989). In *sines plus noise* synthesis, a sound is decomposed into its sinusoidal components (partials) and residual (noise).

In the past decade different improvements to *sines plus noise* synthesis have been achieved, ranging from improved analysis techniques to better understanding of transients. Applications include extracting features from the original sound such as the gender of the person speaking in the case of a voice or the amount of auditory roughness, and the recognition of musical instruments by spectral analysis. Serra's research group, the Music Technology Group of the Pompeu Fabra University,⁴ is a flourishing European computer music centre, involved in technology transfer in

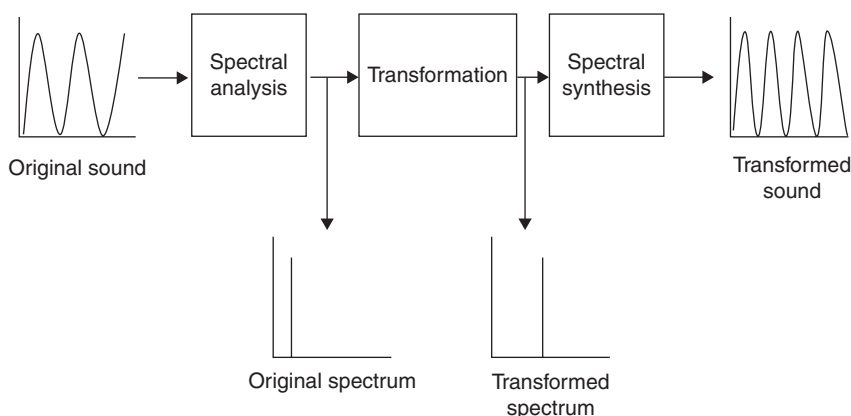


Figure 10.1 A simple analysis–transformation–synthesis representation based on spectral modelling

collaboration with companies, as well as providing free open-source software implementations of much of their work. In some celebrated applications of the sines plus noise model, karaoke demonstrations have been made of live voice transformation where the singer can take on the voice of another (perhaps a famous singer), and through a collaboration with Yamaha, the singing voice synthesiser *Vocaloid*.

Linear predictive coding (LPC) can also be considered as a spectral modelling technique especially useful for voice analysis and synthesis. Originally developed for speech in the late 1960s and early 1970s, it is an example of the technology transfer that tends to occur between computer music and the larger research groups in telecommunications and speech. In LPC the main resonances of a voiced sound are represented in terms of a digital filter; by modifying the parameters of the filters, it is possible to obtain interesting sonic variations. LPC has been extensively used by composers such as Paul Lansky – famously, for *Idle Chatter* (1985) – and Charles Dodge. In his piece *Any Resemblance is Purely Coincidental* (1980), Dodge used a technique of source separation on a 1907 recording of Leoncavallo's aria 'Vesti la giubba' to separate Enrico Caruso's voice from the instrumental accompaniment. Dodge manipulated the voice using LPC, creating new contours and chorus effects.

As with granular synthesis, spectral manipulations of sounds are available in most software synthesisers and are extensively used by composers. Technology has reached a point where sounds can be analysed and resynthesised in realtime easily.

Physical Modelling

Sound synthesis by physical modelling is a class of synthesis techniques in which the source sound production mechanism is mathematically simulated. As opposed to spectral models, physical models do not consider the way the sound is perceived by the ear, but how it is produced by a vibrating object.

In 1961, John Kelly and Carol Lochbaum designed an algorithm to simulate the human vocal tract, considered as a connection of several cylinders with different lengths and widths. This algorithm was used to produce what is perhaps the first musical application of physical modelling synthesis. During a collaboration with Max Mathews, the physical model was used to simulate a human voice with a IBM 704 machine.⁵ Arthur C. Clarke, visiting John Pierce at Bell Labs, heard this demo, and recommended it for the movie *2001: A Space Odyssey*, where the dying HAL9000 computer slowly sings its first song 'Daisy Bell (Bicycle Built for Two)'.

Concerning musical instruments other than the human voice, the first computer simulations by physical models were performed by Hiller and Ruiz in 1971, when a vibrating string was reproduced using numerical methods. To my knowledge such a simulated string was used only for scientific purposes, and no musical compositions were produced with it.

At the end of the 1970s, three acousticians named Michael McIntyre, Robert Schumacher and Jim Woodhouse wrote what is nowadays considered one of the landmark papers on physical modelling synthesis. In 'On the Oscillation of Musical Instruments,' published in the *Journal of the Acoustical Society of America*, the three acousticians described mathematical simulations of three types of instruments: a violin, a clarinet and a flute. Such instruments can be considered as self-sustained oscillators, which means that the sound is produced as long as energy is provided to the system (by bowing or blowing). The paper describes how these three instruments have a very similar algorithmic structure, since they all have a linear element (the vibrating string for the violin or the tube for the flute and clarinet) which is excited by a non-linear element (the bow exciting the string or the player blowing inside the flute).

At approximately the same time Kevin Karplus and Alex Strong developed an algorithm to simulate sounds produced by plucked strings. They noticed that by feeding a circulating buffer with white noise, and adding a low-pass filter at one extremity of the buffer, as shown in Fig. 10.2, it is possible to simulate sonorities similar to those produced by a plucked string. Intuitively, the circulating buffer represents a vibrating string. The shorter the buffer, the higher the frequency of the string. The short noise burst of input simulates the energy imposed to a string at rest when it is plucked into vibration, while the low-pass filter represents propagation energy losses along the string. This simulation is known nowadays as the Karplus–Strong algorithm. The main advantage of this algorithm is its low computational cost.

Julius Smith and David Jaffe extended this algorithm and analysed it from the physical modelling point of view. They improved the excitation and the filters, and added different effects which were not present in the original algorithm. The extended Karplus–Strong algorithm was used by David Jaffe in his piece *Silicon Valley Breakdown* (1982). Jaffe's work demonstrates a number of the compositional possibilities of physical models: the notes are often too fast for any human to play, including intricate time structures like sinusoidal tempo canons; the timbre of the simulated strings can be changed continuously, showing the potential for parameter variation in the model; and imaginary instruments impossible to realise in the real world can be explored, most notoriously, the conceit of plucking a string with the dimensions of the support wires of the Golden Gate Bridge.

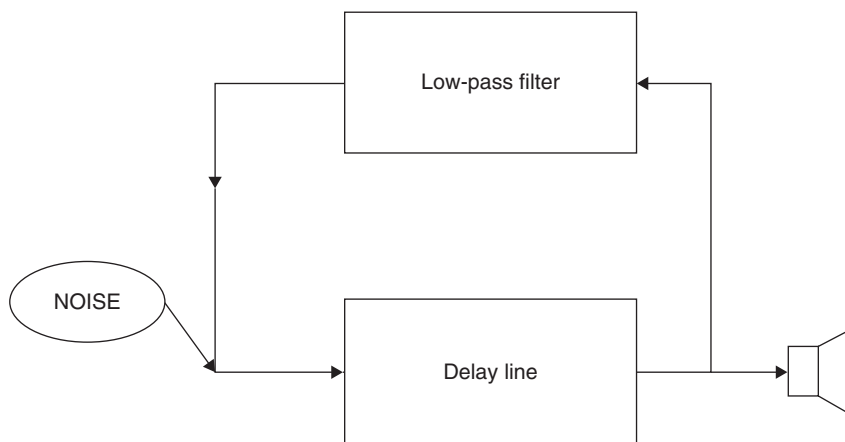


Figure 10.2 The Karplus-Strong algorithm

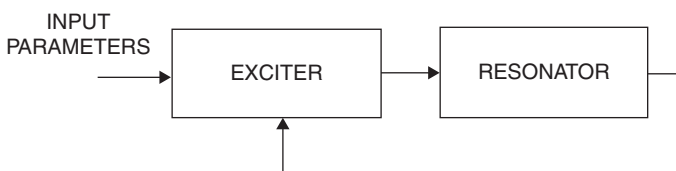


Figure 10.3 The exciter-resonator approach to physical modelling synthesis

In 1990, Charlie Sullivan, then an undergraduate student at Princeton University, developed some extensions to the Karplus-Strong algorithm in order to simulate an electric guitar with distortion and feedback. This improved model was used by composer Paul Lansky in the piece *Things She Carried* (1997).

The extensions to the Karplus-Strong algorithm were part of Smith's development of the digital waveguide theory for physical models. Smith developed a solid physical modelling theory based on the principle of wave propagation in different media. Since then, many musical instruments have been simulated with digital waveguides, which remain the most popular physical modelling synthesis technique. Digital waveguides were also licensed to Yamaha, who in 1994 released the *VLI*, a synthesiser based on physical modelling techniques. Unfortunately, the *VLI* did not have the same commercial success as the *DX7*, being rather expensive, and requiring practice in order to master the different controllers provided with it, such as the breath controller used as an input device to the woodwind physical models.

An interesting aspect of physical modelling synthesis is the decomposition of a vibrating object into exciter and resonator, as shown in Fig. 10.3. Here, the exciter is intended as the source of energy imposed to the

system, while the resonator is the object which produces sound. In Fig. 10.3, exciter and resonators are connected in a feedback loop – this is the case for self-sustained oscillators such as the violin, in which there is a continuous interaction between the bow and the string. In contrast, in percussion instruments the interaction between the player and the instrument is transient for each particular stroke, which means that the player interacts with the instrument for a finite amount of time, and then the instrument is left to resonate. The exciter–resonator approach is particularly interesting from a musical perspective, since unnatural exciters and resonators can be combined together, to create augmented virtual instruments.

Chris Chafe is a composer who has extensively used physical models, and more specifically digital waveguides, in his compositions. In the installation *Ping* (2001), developed by himself together with the digital artist Greg Niemeyer, the Karplus–Strong algorithm again stars. The name *Ping* derives from the Unix command which allows the probing of the distance to a target machine on a network. Chafe decided to sonify such connection time, by considering two locations as if connected by a vibrating string. When the two locations are close together and the network traffic is low, the frequency of the corresponding vibrating string is high. On the other end, when the two locations are far apart or there is a high network traffic between them, the frequency of the corresponding vibrating string is low. In an installation featured at the San Francisco Museum of Modern Art, visitors had the opportunity to choose the different locations they wanted to ‘ping’, resulting in a sonification of the locations. In another attempt to create combinations of exciters and resonators impossible in the real world, Chafe created *Oxygen flute*. *Oxygen flute* is a growth chamber filled with bamboo and carbon dioxide analysers. In this installation, visitors can hear the exchange between their respiration and the respiration of the plants, and a flute physical model is activated by the breathing of the people inside.

A different approach to physical modelling synthesis is *modal synthesis*. Although for a long time very popular in engineering, its introduction to the computer music community is attributed to Jean-Marie Adrien around 1988. Modal synthesis represents a hybrid between physical models and spectral models. A mode, in fact, represents a resonance of a vibrating system, and the modes of each object are a consequence of its physical structure. A synthesiser which implements modal synthesis exclusively is *Modalys*, developed at IRCAM. In *Modalys* the user can choose different resonators and excite them with different inputs. In 1999, Hans Tutschku wrote *Eikasia*, an eight-channel electroacoustic composition, with *Modalys*. While in his previous work the composer had always used pre-recorded sonic material manipulated in different ways, the goal of *Eikasia* was to achieve equivalent sonic complexity

using physical models. The piece uses mostly circular and rectangular plates, whose spectra are tuned according to analysis data of low-frequency piano strings.

A third approach to physical modelling synthesis is represented by *mass-spring models*. In this method, each object is discretised as a finite connection of masses and springs. For example, when simulating a vibrating string a one-dimensional connection of masses and springs is required, while, when simulating a plate, masses and springs are placed in a two-dimensional configuration. Claude Cadoz and his team at the ACROE laboratory in Grenoble are among the pioneers, and developed a software package called Cordis-Anima. In this software, the user can combine different masses and springs to create simulations of existing musical instruments or hybrid objects. An example of a composition written using Cordis-Anima is *pico ... Tera* (2001) by Claude Cadoz himself. The piece uses a single model with thousands of masses and several interacting objects; the five minutes of music in this piece are created by simply running such a model, without any external interaction or post-treatment.

The short history of physical modelling synthesis has shown some successful collaborations between composers and researchers, which have produced both a better understanding of physical modelling techniques and interesting musical compositions. In addition to the examples described so far, a collaboration between composer Juraj Kojis and the author has produced a physical model of a rotating corrugated tube, together with the musical composition *Garden of the Dragon*, in which real and synthetic singing tubes interact. Singing tubes are musical toys, popular in the 1980s, which produce pleasant sonorities when whirled in the air due to regularly spaced corrugations inside. The air travelling within a tube is perturbed by the corrugations, and the frequency of perturbation determines the fundamental frequency. In *Garden of the Dragon*, the performers whirl singing tubes in the air, and such tubes interact in realtime with the virtual tubes simulated in software.

Yet sound synthesis by physical modelling is a technique which often seems more popular among researchers than composers. One of the common criticisms of this technique is the fact that it is pointless to use simulated musical instruments when their real counterparts already hold natural sonic quality. However, aside from the availability of sound models to those who cannot easily obtain and play the real instruments, sound synthesis by physical models becomes musically interesting when sonorities which cannot be achieved with real instruments are produced. When the full potential of physical models is exploited, the composer is able to vary the size and shape of the virtual instruments, or create hybrid connections which are not present in reality.

Overall, physical modelling is much less exploited than other synthesis techniques. The main reason is the fact that fewer software packages which implement sound synthesis by physical modelling are present.⁶ Moreover, physical modelling can also appear frightening to some musicians, since it requires a stronger knowledge of mathematics and physics. This last concern, however, is not always true, since composers using physical models can abstract away from knowing how the model was implemented, and use it as a creative tool controlled by the same parameters as objects in the real world.

The Present and the Future

The availability of software and hardware technology at an affordable price has enormously expanded the quantity of compositions produced using sound synthesis. Sound synthesis is a standard part of curricula and research efforts in many institutions worldwide. Programs like Max/MSP have radically changed the way composers and performers interpret the computer – from being a laborious tool which required lots of time to achieve even a very modest result, the computer has become another musical instrument with which composers and performers can interact in real-time. One aspect which is particularly interesting in the use of interactive sound synthesis programs is the possibility to create interactions between the real and virtual world. Such interactions can take different forms. As an example, augmented instruments use computers as an extension of the possibilities offered by traditional instruments, as discussed in Chapter 5. Lately, augmented instruments have been designed as traditional instruments embedded with sensors. Interaction between the human performer and computer-generated sounds is currently an important topic of research for interaction designers and composers.

In 2001, during a panel on the future of computer music research that took place in Barcelona, Xavier Serra claimed that sound synthesis is dead, since nowadays people are just reinventing the wheel, and no new algorithms on the order of FM synthesis have been invented during the past two decades. Perry Cook, in his book *Real-sound Synthesis for Interactive Applications* (Cook 2002), presents a more optimistic view, claiming that the possibilities offered by sound synthesis are never-ending, since new algorithms and new physical phenomena will always be discovered, which can be applied to sonic simulations.

In a way both Serra and Cook are right. It is true that no profoundly different synthesis techniques have recently been invented. However, researchers are refining existing algorithms to improve aspects such as analysis

techniques or the creation of new sound effects, or are developing novel software platforms. Moreover, composers are using existing algorithms extensively to create art works. Successful traditional musical instruments have a history which spans centuries, and this is obviously not the case for virtual musical instruments. It is important to start thinking about issues such as repertoire and sustainability of virtual musical instruments.

Among the different sound synthesis techniques introduced in this chapter, research on spectral models is currently very active. Spectral analysis techniques are widely adopted in the field of music information retrieval, where analysis is applied to massive databases of audio files. Applications include searching for music by similarities, query by humming, and many more. New hybrid synthesis techniques are also starting to appear, such as concatenative sound synthesis (Schwarz 2004), where a large database of sounds, segmented into units and tagged by sound features, is used as a starting point for producing complex sonic patterns. In the realm of sound modelling, researchers are starting to combine spectral and physical models, to be able to take the advantages of both techniques. This leads to the creation of so-called 'physically informed' techniques, in which spectral data are driven by physical data. To achieve this goal, a better understanding of the relationship between the way a sound is produced and how it is perceived by the human ear is necessary. Hybrid spectral and physical models seem a promising approach which could limit criticisms of physical modelling, since the sound quality of physical modelled sound is often not appreciated by musicians.

From the composers' point of view, however, the main goal in using synthesised sounds will probably always be the potential to create sonorities that do not exist in the real world. As a researcher, I find it rewarding to see composers constantly interested in experimenting with new developments in sound synthesis; a better communication between scientists and composers should be established, since it is rarely the case that one single human being excels both as a researcher and as an artist.

Notes

1 <http://paul.mycpanel.princeton.edu/radiohead.ml.html>

2 Arguably to the point of overuse for certain basic granular techniques in much recent electroacoustic concert music!

3 Research is ongoing; in one recent development, the grains of sounds are analysed for their features and can be used to form large databases of sound material for

compositional reuse. *Concatenative synthesis* (Schwarz 2004) is now a blossoming area of interest for researchers and composers, often bringing up issues of copyright anticipated by the manually constructed works of composers like Pierre Schaeffer or John Oswald.

4 www.mtg.upf.edu

5 Using results from recent research on speech sound by Gunnar Fant.

6 For example, unfortunately, both Modalys and Cordis-Anima have a rather limited number of users, which is probably due to the way the software packages can be obtained.

Modalys needs to be purchased from IRCAM, while, at the time of writing, Cordis-Anima runs only on Silicon Graphics machines, and needs to be purchased from ACROE.