CHAPTER TWO

MUSICAL SOUNDS AND MUSICAL SIGNALS

JOE WOLFE[1]

Music is a type of communication which links the thoughts and feelings of composer, performer and listener, two or all of whom may sometimes be the same person. A musical signal, often transcribed as musical notation on sheet music, is interpreted by a player, who controls an instrument to produce a musical sound, which is transmitted to the ear of the listener, who then decodes the musical message.

Musicians spend years mastering the production and interpretation of musical signals, and the subtle, rapid and coordinated control of musical instruments. However, all artists must understand the fundamental nature of their medium. Music teachers require such an understanding to prepare them for many of the questions that students may ask. Understanding the nature of musical can, at times, help the performer in the eternal quest to produce a better sound and richer performance.

This chapter provides an introduction to musical sounds and signals. Because of its brevity and the depth of the issues raised, an online appendix contains links to further information and sound files. Selected resources introducing the physics and psychoacoustics of music are listed at the end of the chapter.

Sound, vibrations and resonance

The vibrations of a string instrument, the air in a wind instrument, the head of a drum or the cone of a loudspeaker create sound waves that spread out into the air and ultimately produce vibrations in your eardrum, middle ear, and inner ear.

A sound wave is a series of varying displacements of molecules and slight but rapid variations in the pressure of the air (or water or other medium). Air moves from high to low pressure, so a cycle of varying pressure can propagate through the air. The speed of this sound wave is typically 344 metres per second.

[1] Wolfe, J. (2012) "Musical sounds and musical signals" in *Sound Musicianship: Understanding the Crafts of Music*. A. Brown, ed. ISBN 13: 978-1-4438-3912-9.

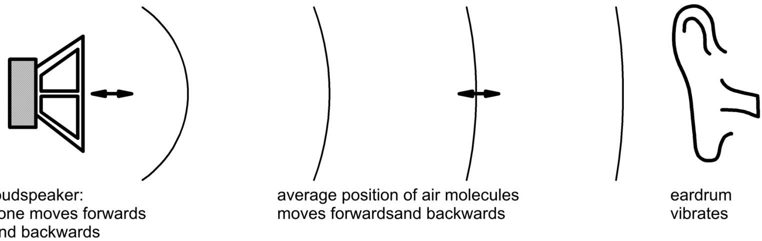

Figure 1. The transfer of vibration from a source, through the air, to the ear, via sound.

The frequency of these vibrations—the number of vibrations occurring in one second—largely determines the pitch. The amplitude of the vibration—millimetres or less for the sources listed above, and perhaps a million times smaller for the vibrations taking place in the eardrum—is important because, combined with the frequency, it determines the loudness.

Some vibrating systems can store energy at a particular frequency—a phenomenon called resonance. A pendulum, for example, stores energy at a single resonant frequency. Strings, columns of air, wooden plates, metal bars and drum skins, in contrast, are all capable of resonant vibrations at several different frequencies.

Frequency and pitch

440 Hz is a standard frequency of the tuning A, called A4 or the A above middle C. From A3 (220 Hz) to A4 (440 Hz) is the interval we call an octave. A4 to A5 (880 Hz) is also an octave, and most observers perceive these musical intervals as equal. It is therefore the ratio of frequencies—not the difference between them—that determines the pitch difference we hear.

Many instruments divide the octave into 12 equal musical intervals, called semitones; this division of the octave is called equal temperament, or ET (strictly 12ET). Be aware that tuning compromises that produce satisfactory thirds and fifths in different keys are called temperaments. The ET semitone is divided into 100 cents. The ET semitone is not the only semitone. On most string and wind instruments (though not pianos, organs or tuned percussion), semitones can be increased or decreased in size during performance; musicians often do so according to other temperaments (discussed later) and musical context.

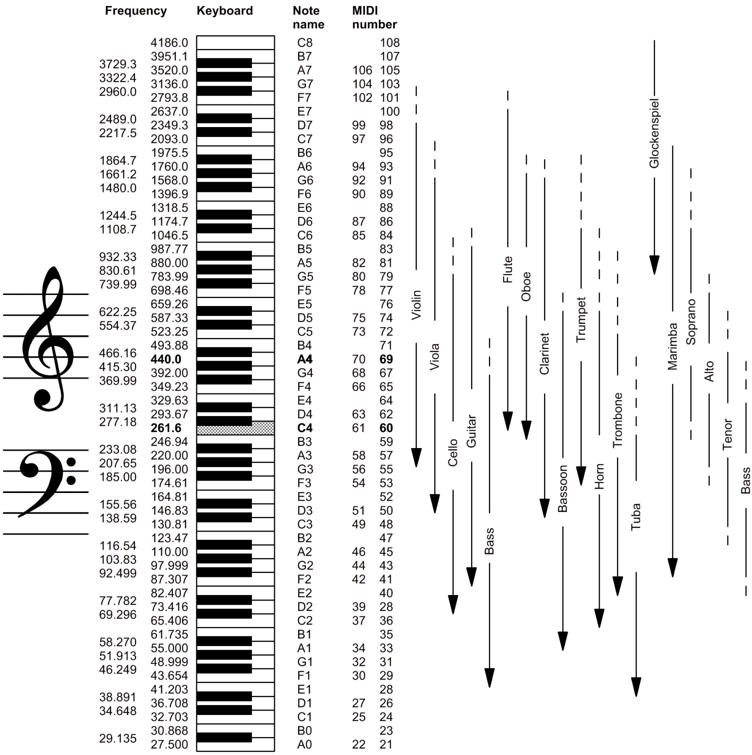

Figure 2. Height on the staff, position on a keyboard, note names and MIDI numbers all relate pitch to the logarithm of the frequency. The range of hearing extends less than an octave below the keyboard, but more than two octaves above.

Intensity and loudness

Loudness depends on the intensity—the power per unit area—at your ear. Intensity is measured in watts per square metre. 1 Wm^{-2} is painfully loud. Under optimal conditions, we can hear sounds a million million times less intense (intensities of only 10^{-12} Wm^{-2}). Sound intensity is proportional to the square of the sound pressure, so we can hear sound pressures over a very wide range.

The decibel (dB) is a logarithmic unit that conveniently reduces this huge range. The sound pressure level in decibels ranges from the reference level of 0 dB (close to the limit of hearing) to 60 dB (conversational speech) up through 120 dB (painfully loud) and beyond.

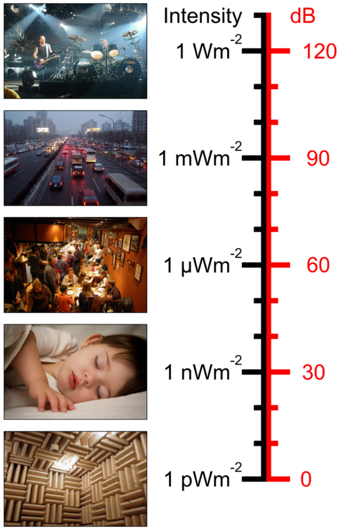

Figure 3. Another logarithmic scale: the decibel (dB) scale relates sound level to the log of sound intensity. The prefixes μ, n and p mean micro (10^{-6}), nano (10^{-9}) and pico (10^{-12}) respectively.

Human hearing is most sensitive in the frequency range 1 to 4 kHz, and much less sensitive at low frequencies, so a tuba player must work harder than a piccolo player to be perceived as equally loud. For this reason, sound level is often measured with the A-weighting filter, which roughly approximates the human hearing curve. Levels filtered in this way are rated in dBA or dB(A). Perceived loudness is difficult to quantify, because it necessarily involves human judgment. Listeners report that a sound is twice as loud if its level is increased by approximately 10 dB.

Doubling the power increases the sound level by 10 log 2 = 3 dB. Increasing the power by a factor of ten increases the level by 10 dB. So, roughly speaking, 2 violins produce 3 dB more sound than one, and 10 violins are 10 dB more than (or twice as loud as) one.

At close range, some acoustic instruments can produce sound levels of over 100 dB. Exposure to such levels can lead to hearing loss and is a serious occupational hazard for orchestral as well as rock musicians. Serious hearing problems can also arise as a long-term result of loud music played on personal music systems.

Pure tones, harmonics, periodic and non-periodic sounds

A tuning fork produces a pure tone: that is, it produces a sound pressure that varies sinusoidally with time. A more complex sound may be considered as the sum of many sine waves; the sound spectrum is like a "recipe" for creating that sound. Each sine wave component in the spectrum has its own amplitude, which expresses how much of each sine wave partial is present, as illustrated in Fig. 4.

A harmonic spectrum has components in the harmonic series, which has important consequences for music. If we start with a pure tone with the fundamental frequency f_1 (for example, take the note at A2, the A near the bottom of the bass clef at 110 Hz), then the harmonic series contains these frequencies:

Generally: f_1, $2f_1$, $3f_1$, $4f_1$, $5f_1$… nf_1 Hz, where n is an integer.
Example: 110, 220, 330, 440, 550… n×110 Hz

The above series is familiar to most musicians. For example, a string player can play the harmonic series as "natural harmonics" where the open string gives f_1, touching the string halfway along gives $2f_1$, touching the string at one third of its length (a "touch fifth") gives $3f_1$, at one quarter (a "touch fourth") gives $4f_1$, and at 1/n of its length gives nf_1. Brass players, and some woodwind players, can usually play several of the notes in this series without moving the instrument's valves or keys.

Something interesting happens when we add tones in the harmonic series (Fig. 4). When the first component (at f_1) has finished one complete vibration, the second has completed exactly two vibrations, the third exactly three vibrations and the nth exactly n vibrations. So, after one vibration of f1, all the harmonic vibrations have completed a whole number of cycles. Consequently, their sum for the next cycle looks exactly the same. This makes it a periodic vibration: after one cycle of the fundamental, another identical cycle is produced. In our example above, the cycle lasts (1/110) seconds = 9 milliseconds or 9 ms, which we call the period, $T = 1/f_1$.

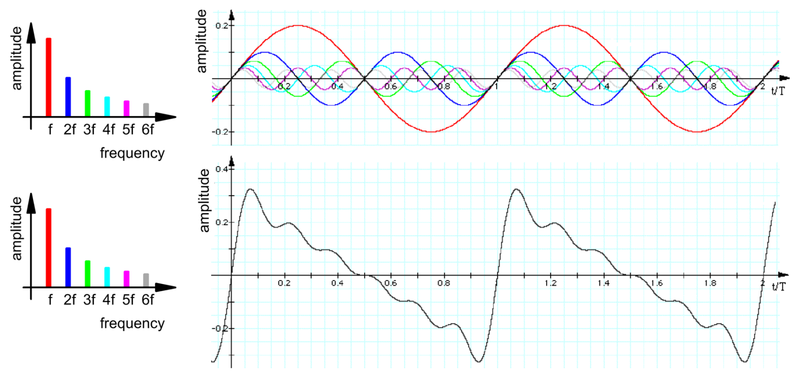

Figure 4. Each diagram at left shows a harmonic spectrum. The bar at each frequency represents the amplitude of the sine wave at that frequency, plotted at top right. After one cycle of the lowest frequency, all of the vibrations start again, so the total vibration repeats itself exactly. This is a periodic vibration, which has a clear pitch.

If the higher frequency components are strong, this periodic vibration or harmonic spectrum sounds like a chord. If they are weak, it sounds like a note with the same pitch as a pure tone with fundamental frequency f_1. If the fundamental is removed, the combined wave still has a period $1/f_1$, and still has the same pitch, though it sounds less "bassy". This is why a tiny loudspeaker or a telephone—both of which transmit negligible power at low frequencies—can still enable us to hear the pitch of bass instruments.

A harmonic spectrum therefore implies a periodic vibration. Conversely, a periodic sound has a harmonic spectrum. The human voice, bowed strings and wind instruments produce almost exactly periodic sounds, with rare exceptions such as multiphonics.

Musical instruments have resonances: strings and air columns can store energy in vibrations at particular frequencies. An ideal string (uniform and infinitely flexible) has harmonic resonances, as described above, and so creates harmonic spectra when struck. A real string (especially if thick and/or made of steel) cannot bend sharply at the bridge, and consequently its resonances are spaced a little more widely than the harmonic series. Wound strings reduce this problem: the winding provides the required mass, while the thin core allows sharp bends. Slightly non-ideal strings cause octaves at the extremes of the range of pianos to be slightly stretched: the tuner tunes a string to the second resonance of another, nominally an octave below. Guitar strings become less ideal as they wear over the frets; they will eventually require replacing as they become less harmonic and harder to tune.

The strings of bowed stringed instruments also have slightly inharmonic resonances. A rosined bow drives a string by "sticking" to it very briefly and dragging the point of contact slightly sideways, after which the string "slips" back in the opposite direction. This produces periodic cycles of vibration—they repeat each cycle exactly, as in Fig. 4—even if the frequency components are thus slightly displaced from the resonances. Strings become non-ideal as they accumulate rosin or grease, and may require cleaning.

The resonances of woodwind and brass instruments are typically less perfectly harmonic than strings, and those of the vocal tract are not harmonic at all. Usually, the vibrations of the vocal folds, an air jet, a reed or a player's lips are periodic. In normal performance, these instruments thus possess harmonic spectra. In both wind instruments and bowed stringed instruments, the production of harmonic spectra is technically known as nonlinear behaviour, while a plucked string exhibits linear behaviour.

This nonlinearity has important applications: until the technology to make uniform strings became available, only nonlinear instruments, such as the human voice and early wind instruments, produced the harmonic spectra that are fundamental to our current idea of harmony, as shown below.

With tuned percussion instruments, few or none of the higher resonances are tuned to an almost harmonic relationship with the first. The bars of xylophones and marimbas have an arch cut into the bottom so that the frequency of the second resonance falls close to three and four times that of the first, respectively. The first resonance determines the pitch in these cases, but the short-lived, inharmonic higher-frequency components are vital to the percussive sound. In the case of timpani, the frequency components fall near 2f and 3f. We hear a pitch corresponding to f—which, as it is not present in the spectrum, is an example of the missing fundamental.

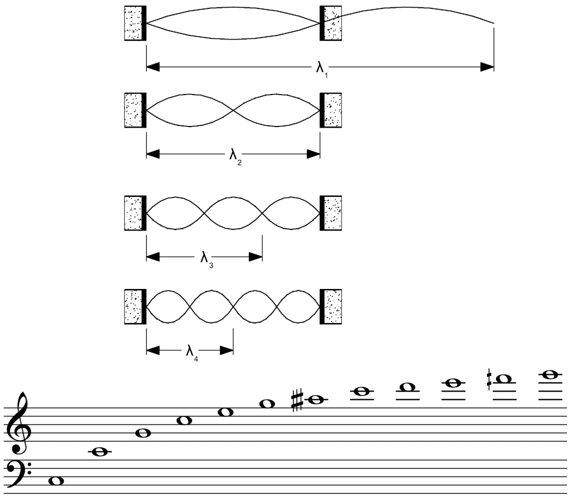

Figure 5. The figure shows the first four resonances of a uniform, stretched string of length L. The wavelengths λ of standing wave on the string are shown.[2]

In untuned percussion (triangle, drums etc.), the spectrum is very inharmonic and we hear no clear pitch, although we can tell that a tenor drum is higher than a bass drum. Cymbals, gongs and tam-tams have the interesting (nonlinear) property of transferring energy from high to low or low to high frequencies during a note, depending on the shape of the instrument.

Timbre, spectrum, envelope and transients

Timbre is defined negatively: if two sounds have the same pitch and loudness, yet are perceived as different, the component that makes them subjectively distinct is the timbre. Timbre depends on the spectrum of a note, and especially on how the spectrum varies over time.

---

2. If the speed of the wave in the string is v, then the frequencies f of the resonances are $f_1 = v/\lambda_1 = v/2L$, $f_2 = v/\lambda_2 = 2v/2L$, $f_3 = v/\lambda_3 = 3v/2L$, etc., giving the harmonic series f_1, $2f_1$, $3f_1$, etc. By lightly placing a finger at one of the nodes, players can excite what they call natural harmonics with fundamentals in this series. The notes show the first 12 notes in the harmonic series, beginning on C3 in this case. The 7th and 11th notes fall closer to a quarter tone than they do to notes in the chromatic scale.

In a harmonic spectrum such as in Fig. 4, the relative strength of harmonics contributes to timbre: a note with stronger high-frequency components has a bright timbre. For reasons related to the nonlinearities mentioned above for the voice and wind instruments, high harmonics rapidly become stronger as the loudness increases. Playing loudly consequently produces a brighter sound, which can be distinguished as loud playing even when a recording is replayed at low volume.

The overall shape or envelope of the spectrum can be altered by mutes on brass instruments, effects pedals on electric guitars, and filtering in the studio. Broad peaks in the spectral envelope are called formants, and their frequencies are important to the timbre of instruments and to speech sounds.

Whether or not a spectrum is harmonic makes a large difference to the timbre: for example, the inharmonicity of higher harmonics is important to the timbre of a plucked string and vital to that of a bell.

The variation of a note's properties of a note over time is more important to recognition than is the spectrum. For instance, a bowed violin note without vibrato varies little over time, while a pizzicato note becomes rapidly less loud and its higher frequencies fade rapidly. As a result, a recording of a pizzicato or piano note is completely unrecognisable when played backwards, sounding a little like a badly adjusted harmonium.

Vibrato—a cyclic variation in the note with several cycles per second—also contributes strongly to timbre and expressiveness. For the voice, vibrato is mainly a cyclic variation in pitch. On many wind instruments, vibrato also involves a strong variation in loudness as the player varies air pressure in the mouth and airflow into the instrument. (Cyclic variation in loudness alone is called tremolo.) On instruments like violins the cyclic rocking of a finger on the fingerboard produces a cyclic change in pitch. However, because of the dramatic way in which the output of a violin depends on frequency, this modest variation in pitch produces a strong cyclic variation in spectrum that is essential to the violin's sound.

Perhaps the most important indentifying parts of a note's timbre are the initial and final transients. During the first tenth of a second or so, the spectrum contains relatively high proportions of inharmonic components which are often vital to identifying an instrument.

Harmony, scales and temperaments

Harmonia (αρμονια) is Greek for "fitting together". The reasons why some notes seem to fit together involve physics, physiology and fashion.

First, consider what happens when we add two tones. If their frequencies are exactly the same, they are heard as a single tone. If the frequencies differ by up to several Hertz, we hear interference beats: in other words, a tone with intermediate frequency whose loudness varies cyclically. Musicians often tune up by slowing down, and then removing, the beats created between their instrument and another. The number of beats per second equals the difference in frequency between the two tones. Slow beats are part of the appeal of the "chorus" effect generated when, for example, a section of violins plays together.

If we gradually increase the difference in frequency, we hear faster beats. Then, somewhere near the point at which separate beats are no longer audible, many people hear a rough sound. At a larger frequency difference, we start to hear two separate pitches. If the two tones are sufficiently high, sufficiently separated and sufficiently loud, we can often hear a note whose frequency equals the difference between those of the two tones—this is called the Tartini tone. (See the online appendix for sound examples.)

The importance of the octave is recognised by our note names: A3 lies an octave above A2, yet both are called A. If a woman sings a phrase, then a man is asked to repeat it, he will usually sing it an octave lower without thinking that he has transposed it. When men and women sing together in octaves, we usually don't regard this as harmony.

There is a physical reason why the octave is special: consider again the harmonics of the note A2, along with those of A3.

A2	110	220	330	440	550	660 Hz, etc.
A3		220		440		660 Hz, etc.

The harmonics of the higher note are a subset of those of the note an octave lower; when we add the higher note, we add no new frequencies.

In many cultures, the fifth (ratio 3:2) is regarded as the next most consonant interval. If we add E3 at 110 Hz×3/2, their harmonics are

A2	110		220	330	440		550	660 Hz, etc.
E3		165		330		495		660 Hz, etc.

Here, every second harmonic of the higher note coincides with one of those of the note a fifth below, and the other harmonics are sufficiently well-separated to avoid the roughness that can be caused by closely grouped harmonics.

In Western music for the last several centuries, major and minor thirds (ratios of about 5:4 and 6:5) and sixths (ratios of about 5:3 and 8:5) are also regarded as consonant: these can be combined to form major and minor chords in which all notes are consonant.

The fact that harmonious musical intervals and string lengths were related in ratios of small integers was known to Pythagoras (sixth century BC), and is one of the earliest examples of a quantitative scientific observation based on the relation between numbers and the natural world.

The major scale (in just intonation tuning) has notes with the following frequency ratios:

	do		re		mi		fa		so		la		ti		do ...
ratio to root	1		9/8		5/4		4/3		3/2		5/3		15/8		2
step ratio		9/8		10/9		16/15		9/8		10/9		9/8		16/15	

Note that the triads on do, fa and so include the ratios 5/4 and 3/2. These chords sound very pure and consonant. (The triad on re, however, sounds very rough.)

Making the approximation that $9/8 \cong 10/9$ (they actually differ by 1%), we have a diatonic scale (five big steps and two small ones). Dividing each larger interval into two gives 12 steps to the octave.

Intervals on an equal-tempered keyboard only approximate the pure ratios discussed above. The ET (equal-tempered) fifth (7 semitones) comes close: $2^{7/12} = 1.498 \cong 3/2 = 1.500$: a difference of 0.1%, 2 cents or 2% of a semitone. The fourth ($4/3 = 2/(3/2)$) is equally close. While these ET intervals produce audible beats, they raise only minor problems. The ET major third is $2^{4/12} = 2^{1/3} = 1.260$. This is larger than the pure third (1.250) by 0.8% or 14 cents, which causes practical problems.

Imagine two flutes are sounding the ET major third A5 and C#6 (880 Hz and 1109 Hz). We hear a difference tone of 229 Hz. Compared with the A3 (220 Hz) being played by the clarinet behind them, this is 67 cents sharp: it is closer to A# than to A. Although the musicians could ignore it, this would be rather uncomfortable. Alternatively, they could use wide vibrato to hide the problem. Often, however, Flute 1 will lower the C#6 from an ET to a just third above the A6. This brings the Tartini (difference) tone in tune, removing the beating—a useful tuning trick.

Pianos cannot do this. However, the three strings producing one note on a piano are already each tuned to a slightly different frequency. This not only increases the sustain time of the note, but also provides a complex set of beats that masks the interference from equal-tempered triads.

Tuning a keyboard using only pure intervals runs into problems, especially if the music is played on many keys. The problem is that F# and G♭ are usually played as the same note, as are E# and F, etc. We can construct a circle of 12 ascending fifths as follows: C-G-D-A-E-B-F#-C#-G#-D#-A#-E#-B#. If we set B# = C, we have covered (exactly) seven octaves. However, while seven octaves = $2^7 = 128$, 12 pure fifths = $(3/2)^7 = 129.75$, which exceeds 128 by 1.4%, or a quarter of a semitone.

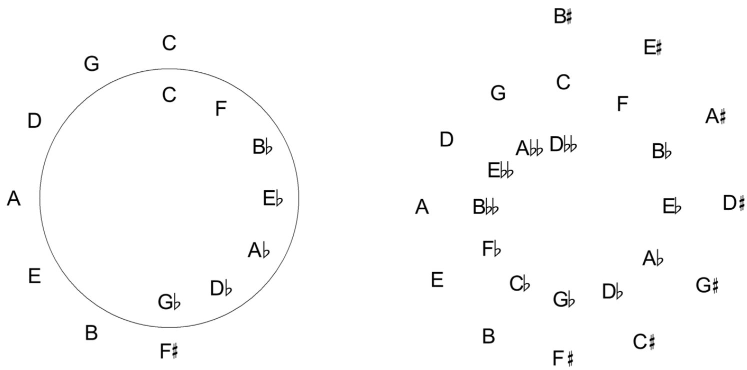

Figure 6. At left, the circle of fourths (clockwise) or fifths (anticlockwise), which implies a temperament that "closes the circle". At right, the spiral created by pure fourths or fifths.

Further, the major thirds that result from tuning pure fifths have a ratio of 81/64, which is even sharper than the ET third.

Music as a signal

Signals are often divided into analogue and digital. An analogue quantity varies smoothly. Examples include the position of a vibrating violin bridge or loudspeaker cone, the pressure in a sound wave, and the voltage produced by a microphone. In modern recordings, microphone voltages are digitised such that at regular time intervals (often 44,100 times per second) the voltage is recorded as a number. Compared to analogue signals, this set of numbers—a digital signal—is less susceptible to distortion and noise, and may be copied many times from one medium to another without any loss in quality. It can also be processed rapidly. When it is converted back to an analogue signal, the discontinuities between voltage values are smoothed out and the resulting signal is often amplified and delivered to speakers, from which it travels to our ears as an analogue sound wave.

Standard music notation is a digital signal. Pitches have particular values and notes have lengths, both of which can be expressed as numbers (and can thus be stored in note processors such as Sibelius). This coding is extremely efficient, with the note processor file of a score being much smaller than the file of a recorded performance. Given the ease and speed with which we can read, process and produce music, it is almost certain that we can remember and process it as a digital signal. See Wolfe (2003) and Wolfe and Schubert (2010) for more detail.

Further reading

The way in which music is produced by different instruments is a fascinating but broad topic, covered by many authors. Sundberg (1989) and Rossing (1982) offer relatively easy introductions, while Fletcher and Rossing (1998) and Benade (1976) are more technical. An online introduction to music acoustics is available at http://www.phys.unsw.edu.au/music. An online appendix to this paper, providing both sound files and more information, is available at http://www.phys.unsw.edu.au/jw/musical-sounds-musical-instruments.html.

References

Benade, Arthur H. 1976. *Fundamentals of Musical Acoustics*. New York: Oxford University Press.

Fletcher, Neville H. and Thomas D. Rossing. 1998. *The Physics of Musical Instruments*. New York: Springer-Verlag.

Howard, David M. and Jamie Angus. 2006. *Acoustics and Psychoacoustics*. Oxford: Focal Press.

Rossing, Thomas D. 1982. *The Science of Sound*. Reading, Mass.: Addison-Wesley.

Sundberg, Johan. 1989. *The Science of Musical Sounds*. San Diego: Academic Press.

Wolfe, Joe. 2003. "From Ideas to Acoustics and Back Again: The Creation and Analysis of Information in Music." In *Proceedings of the Eighth Western Pacific Acoustics Conference*. Melbourne: Australian Acoustical Society.

Wolfe, Joe and Emery Schubert. 2010. "Did Non-vocal Instrument Characteristics Influence Modern Singing?" *Musica Humana* 2:21–38