

---

# Notes on the Theory of Sound

---

This chapter presents concepts that are indispensable to understanding and studying the phenomena associated with sound. In this study, readers can find the details required to discuss the variety of ways that exist to process sounds.

Mathematical equations are deliberately presented in their simplest and purest forms, without going into detail and avoiding any mathematical proofs. However, the standards of scientific rigor to which all physical sciences are committed have not been lost.

## 1.1. Basic concepts

We shall begin by describing the nature of a sound, followed by a few of its characteristics, before turning to the question of how our ears work. This leads into an analysis of the *typology of sounds*, *spectral analysis* and *timbre*.

To conclude, we will present the fundamental aspects of the propagation of sound, and we will consider a few common phenomena to be explored in more detail in subsequent chapters.

### 1.1.1. What is sound?

Although this question might seem simple, the answer is by no means easy. There are two ways to approach this topic: from a purely scientific perspective,

working from the laws of physics, or alternatively by thinking about how our senses allow us to perceive sound.

Physicists view sound as a *mechanical wave* that propagates as a perturbation within an elastic medium or object. In other words, they view sound as the forward-and-backward motion (mechanical oscillation) of particles around their resting position. Electromagnetic waves, on the other hand, propagate as energy in the form of an electric field coupled to a magnetic field.



**Figure 1.1.** *A simple example of a mechanical wave. Here, the wave is created on the surface of the water after a stone is thrown in*

Many people will find it easier to define sounds as auditory sensations.

Sounds are produced by vibrating objects. These objects are sources, and the environment in which the sound is emitted carries the sound to our ears. When a sound reaches our ears, our brains allow us to perceive it, become aware of it and interpret it.

Most of the objects around us can produce sounds when they interact with shocks, friction, airflows or deformations. Who has not entertained themselves by twanging a plastic ruler at the edge of a table?



**Figure 1.2.** *Vibrating ruler at the edge of a table*

Sound cannot propagate in a vacuum, since there is no medium to convey the vibration.

Like all physical phenomena, sounds can be characterized. We can use parameters such as the *intensity*, the *pitch* and the *timbre* to define and distinguish different sounds. Other parameters are possible too – people interpret sounds as they hear them. This involves subjective phenomena, so-called *psychoacoustic* phenomena, which depend on the physiology, culture and ethnicity of the individuals receiving the sound message. This approach to decrypting the process of sound characterization is highly complex. We will quickly encounter questions to which science has not yet been able to provide comprehensive answers.

Over the centuries, philosophers have often wondered: does a sound exist if no one is there to hear it?

The field of science that studies the physics of sound phenomena is known as *acoustics*. Some of its goals include characterizing the *audibility* of sounds, defining

the means of transportation and transformation of sounds and determining the deformations that a given sound can undergo. It is often an extremely theoretical field. Acoustics provides working foundations, but quickly reaches its limits when sounds come together to form music. Music is not a science, and appeals to many parameters that are not necessarily measurable, and which are intricately intertwined in extremely complex ways.

Before we continue, we shall review a few concepts that are required to understand the foundations and the nature of sound.

### 1.1.2. *Intensity*

This parameter characterizes the strength of a sound, determining whether the sound seems loud or quiet. The term *loudness* is also used.

When a sound is emitted, the sound waves deform the fluid medium (usually air). This deformation creates a change or local perturbation in the pressure around the source of the sound. This perturbation travels through the surrounding material(s) at a speed (*speed of sound* or *celerity*) that depends on the nature of the elements in the medium, as well as their states and thermodynamic properties. The materials through which the perturbation travels – fluids or alternatively flexible or rigid bodies – each have a certain *elasticity*, which usually allows them to recover their original states once the sound wave(s) have passed through them. Permanent deformation or destruction can occur if the sound pressure generated by the wave is greater than the *elastic limits* of these materials. This scenario is uncommon, since the stresses involved in sound propagation are relatively low.

The pressure variation mentioned above, usually observed in air, is known as the instantaneous acoustic pressure, measured in *pascals*. This acoustic pressure induces an acoustic energy. These two parameters are both described by the term *acoustic sound pressure level*, also simply called the *sound level*.

The sound pressure ranges over many different scales. Its units are *pascals* ( $1 \text{ Pa} = 1 \text{ N/m}^2$ ), denoted by the symbol Pa. The normal atmospheric pressure at sea level is defined as  $1.013 \times 10^5$  or 1,013 hPa (1 hPa = 100 Pa). When working with the scale of the acoustic sound pressure, we normalize by a reference pressure value close to the average absolute intensity threshold of the human auditory system between 1,000 and 4,000 Hz (hertz<sup>1</sup>), namely  $2 \times 10^{-5}$  Pa. This corresponds to a

---

<sup>1</sup> Unit measuring the frequency of a phenomenon that has a period of 1 s, see section 1.1.3 of this chapter.

power of  $10^{-12}$  W (watt<sup>2</sup>)/m<sup>2</sup>. However, these numbers are not easy to manipulate and are not very “intuitive”, not to mention the fact that the ratio between the lowest audible sounds and extremely loud sounds is 1/10,000,000.

The sensation of sound is first and foremost a physiological phenomenon, and one useful approach is to quantify the sensation of excitation in a way that takes into account the acoustic range of the human ear, which works according to a logarithmic scale. The inventor of the telephone, Graham Bell<sup>3</sup>, originally defined the *bel* as a base unit. Today, we often use the unit of one-tenth of a bel, the *decibel* (dB), as a simple way to quantify sound phenomena.

$$N_{dB} = 20 \log_{10} \frac{P}{P_0} = 20 \log_{10} \frac{P}{10^{-5}}$$

or

$$N_{dB} = 10 \log_{10} \frac{W}{W_0} = 10 \log_{10} \frac{W}{1 \times 10^{-12}}$$

where:

- $N$ : sound level in decibels;
- $P$ : measured acoustic pressure;
- $P_0$ : reference acoustic pressure (0 dB);
- $W$ : measured power;
- $W_0$ : reference power.

The logarithmic growth of the decibel scale means that the sound intensity doubles every 3 dB.

---

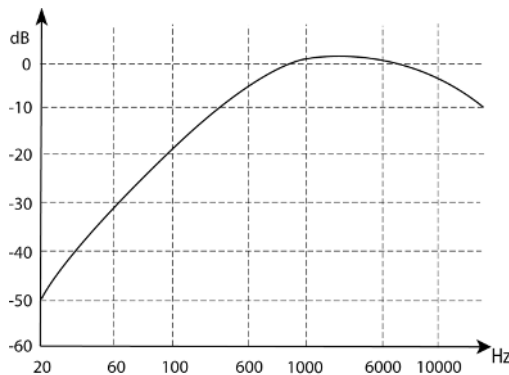
<sup>2</sup> Unit of power. One watt is the power of a system that receives a constant energy flux of 1 J/s.

<sup>3</sup> Alexander Graham Bell, March 3 1847–February 8 1922, American inventor originally from Great Britain, born in Edinburgh. He invented the telephone in 1876 and founded the telephone company bearing his name. Creator of the “National Geographic Society”.

Type of audio signal	Effects	Sound level (dBA)
Rocket take-off		180
Turbojet engine, airplane take-off		140
Rifle shot, engine on a test bench		130
Formula 1, jackhammer	Pain threshold	120
Rock band, metal workshop		110
Train passing by, circular saw, night club		100
Portable music player at maximum volume, sander, shouting	Hearing risk threshold	90
Radio at maximum volume, machine tools		80
Noisy restaurant, office with typewriters	Office work	70
Lively conversation, street, public place		60
Quiet conversation, large quiet office	Intellectual work requiring high concentration	50
Quiet apartment, quiet office		40
Walk through the forest		30
Peaceful countryside, whispering		20
Recording studio		10
Silence	Audibility threshold	0

**Table 1.1.** Table of sound intensities

When measuring the sound level with a sonometer, our unit of choice is the *decibel A* or  $dB(A)$ . As we will see in section 1.2.2, the sensitivity of the human ear varies as a function of the frequency of the sound signal. To compensate for this physiological behavior of our ears, the sound levels of each frequency component of a sound wave are weighted and summed to give an overall measurement. The units of  $dB(A)$  are adjusted to reflect this weighting.



**Figure 1.3.** Psophometric curve (weighted curve)  $dB(A)$

### 1.1.3. Sound pitch

The pitch of a sound is characterized by its *frequency*, i.e. the number of oscillations per second of the molecules in the traversed medium (usually air) around their resting position when a sound wave passes through them.

The frequency is measured in units of *hertz*, denoted by the symbol Hz, and its multiples: kilohertz (kHz), megahertz (MHz), etc. The range of audible frequencies for humans is 20–20,000 Hz (20 kHz). This range varies from individual to individual, and also changes with age.

The speed at which a sound wave travels, also known as its celerity, is 343 m/s through air at a temperature of 20 °C. This value changes depending on the nature of the object or the medium through which the wave is traveling, as well as the pressure and the temperature. For instance, it is equal to 331 m/s through air at 0 °C. The speed of sound is higher in liquid and solid objects (~1,400 m/s in water and ~5,000 m/s in steel).

Having defined the frequency, we should take the opportunity to define a few other parameters: the *wavelength*, the *period* and the *amplitude*.

The wavelength is the distance between two consecutive maxima of a sound wave. This defines the separation between two consecutive periods of a periodic wave (see Figure 1.4), and so is also equal to the distance traveled by the wave during one period.

The period is the time in seconds taken to complete one full oscillation (one cycle). It is the inverse of the frequency and vice versa:

$$T = \frac{1}{f} \text{ and } f = \frac{1}{T}$$

where:

- $T$ : period in seconds;
- $f$ : frequency in hertz.

A frequency of 1,000 Hz (1 kHz) corresponds to a period of 0.001 s or 1 ms.

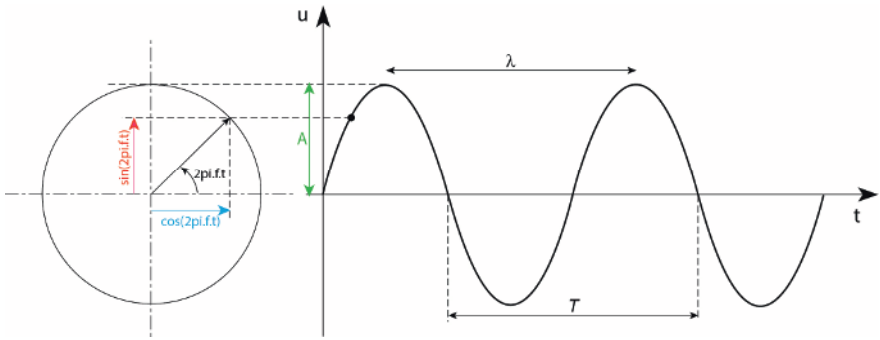
The amplitude defines the sound intensity. As we saw above, this characterizes the variation in pressure.

The wavelength, the period, the frequency and the speed of sound satisfy the relation:

$$\lambda = cT = \frac{c}{f}$$

where:

- $\lambda$ : wavelength in meters;
- $c$ : speed of sound in m/s;
- $T$ : period in seconds;
- $f$ : frequency in hertz.



**Figure 1.4.** Representation of a sound wave and its parameters over time

### 1.1.4. Approaching the concept of timbre

Timbre is what allows our ears to distinguish and recognize different sounds, whether everyday noises, musical instrument or people's voices.

It is closely related to the shape of the sound wave. Timbre is a complex notion that we still struggle to fully explain today, since it involves concepts relating to the act of hearing, our ability to judge sounds, our auditory memory and perception.

Before we can begin to discuss the notion of timbre more precisely, we will need to study several other sound-related parameters, including the physiological mechanisms of our ears, the typology of sounds, the spectrum, transient phenomena and the nature of sound-emitting source(s).

Claude Elwood Shannon<sup>4</sup>, a renowned mathematician, once stated: “Timbre is what makes sounds sing to our ears”.

## 1.2. The ears

Hearing is the second of our five senses. It relies in part on our auditory system, whose primary external organ is our ears.

In this chapter, we will find out precisely how our ears work, what makes them interesting and we will analyze how they operate within sound-based environments.

### 1.2.1. How our ears work

The ear can be divided into three parts: the outer ear, the middle ear and the inner ear.

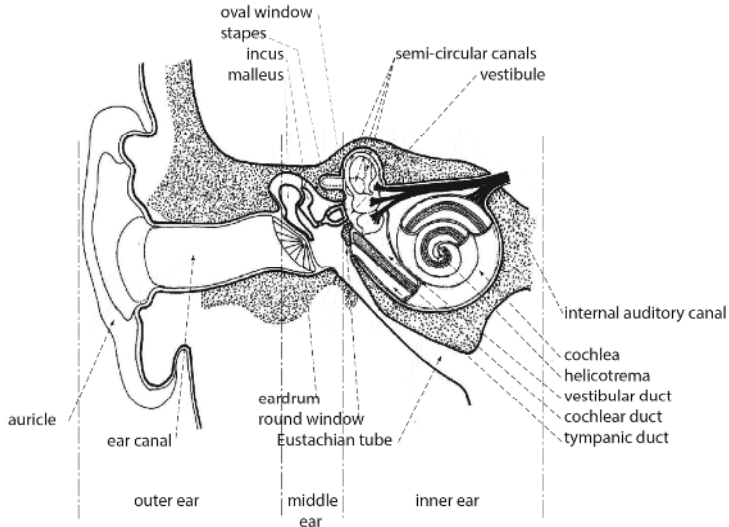
The outer ear is the part of the system that captures sounds. This system serves the roles of amplification and protection. It is separated from the middle ear by a thin, flexible membrane, the eardrum, which deforms under the effect of sound waves.

The outer ear consists of the *auricle* and the *ear canal*, the latter of which is approximately 2.5 cm long. The ear canal carries sound vibrations to the *eardrum*, amplifying frequencies between 1,500 and 7,000 Hz by 10–15 dB. These are the most useful frequencies to us, notably including speech.

The auricle also plays a role in locating the source of sounds in space for frequencies between 2,000 and 7,000 Hz.

---

<sup>4</sup> Claude Elwood Shannon, 1916–2001, American electrical engineer, mathematician and cryptographer, known as the father of the digital transmission of information. He created the field of “mathematical information theory”.



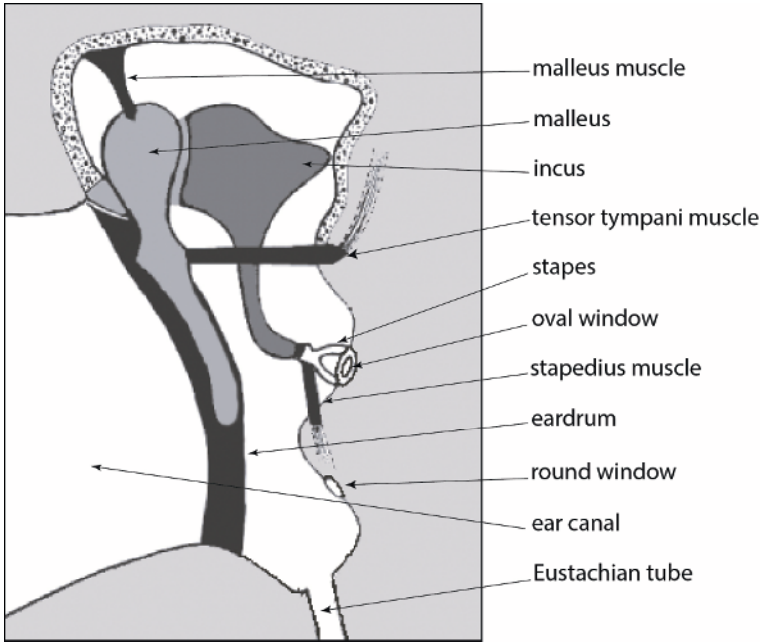
**Figure 1.5.** *The outer, middle and inner ear*

The *middle ear* is a cavity filled with air. It is connected to the *pharynx* by the *eustachian tube*, which opens when we swallow in order to equalize the pressure on either side of the eardrum. This cavity also contains a *series of ossicles* (small bones): the *malleus*, attached to the eardrum, the *incus* and the *stapes*. The stapes acts as an interface between the air-filled medium of the middle ear and the liquid medium of the inner ear. It rests against the *oval window*, which acts as the boundary to the *inner ear*.

Together, the ossicles form a lever that increases the pressure and thus the amplitude of sound waves. The surface area of the eardrum is around 15 times larger than the oval window, which creates an increase of 20 dB. The middle ear acts as a pressure amplifier.

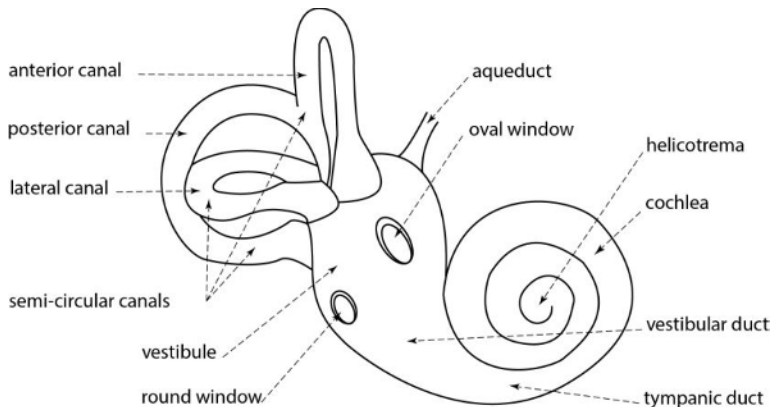
When excessive sounds louder than 80 dB are detected, the *stapedius* (stapes muscle) contracts to reduce the vibration of the ossicles (*acoustic reflex*), attenuating the transmission of sound waves to the inner ear. This protection mechanism reduces the sound signal by 40 dB.

It should be noted that the stapedius develops fatigue over time, and so cannot provide long-term protection. Additionally, it only activates at low frequencies of less than 1 kHz, and the contraction occurs too slowly to protect from sudden noises like explosions, since the “reflex latency” (physiological reaction time associated with the information processing sequence of humans) is around 30 ms.



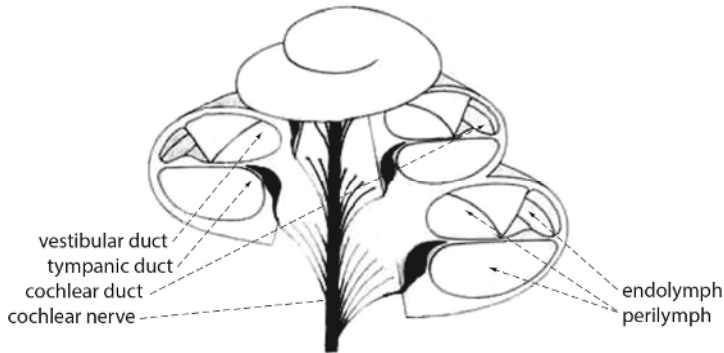
**Figure 1.6.** Detailed diagram of the middle ear

The inner ear consists of two sensory organs: the *vestibule*, a balance organ, and the *cochlea*, a hearing organ.



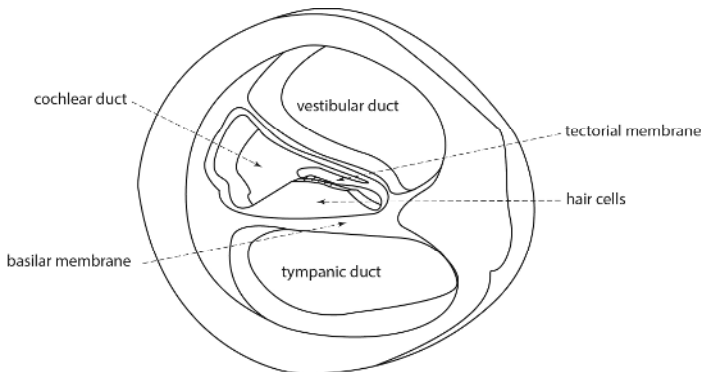
**Figure 1.7.** The inner ear

The cochlea is a hollow bone shaped like a snail with a 2.5-turn spiral. It has three canals: the *vestibular duct*, the *tympanic duct* and, in the center, the *cochlear duct*.



**Figure 1.8.** *The cochlea*

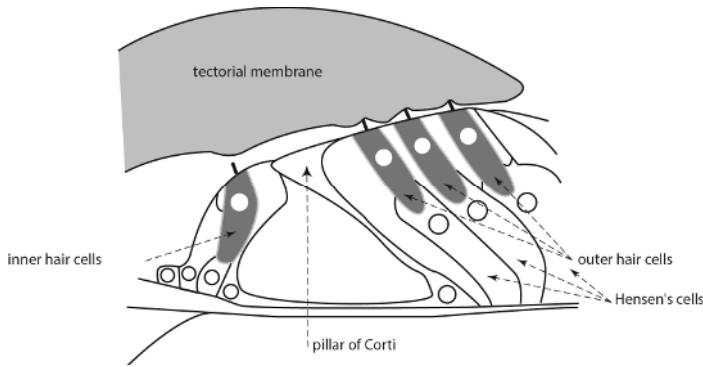
The vibrations applied to the oval window by the stapes are transmitted via the oval window into a sodium-rich liquid called the *perilymph* inside the vestibular duct, which is connected to the *apex* (tip) of the cochlea, in a region called the *helicotrema*. The round window in the tympanic duct compensates for the expansion of the liquid.



**Figure 1.9.** *Cross-section of the cochlea*

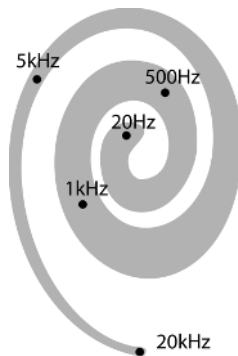
The cochlear duct is closed and filled with potassium-rich *endolymph*. It is separated from the tympanic duct by the *basilar membrane*. The sensory organ carrying the receptors, called the *organ of Corti*, is located on the internal face of

this membrane. The receptors are *hair cells*. The inner hair cells are arranged in a single row, and the outer cells are arranged in three rows. The tips of the hairs (or *stereocilia*) of the outer cells are anchored to the *tectorial membrane*. Both inner and outer cells are connected to the fibers of the *auditory nerve*. The motion of the liquids contained in the ducts induces deformations in the basilar membrane, which tilts and twists the hairs connected to the tectorial membrane, which itself remains fixed. This tilting and twisting encodes the sound vibrations into ionic motion that polarizes or depolarizes the membrane of each cell.



**Figure 1.10.** Detailed diagram of the organ of Corti

The outer hair cells are in fact selective amplifiers, and the inner cells are the actual sensory cells. A *tonotopy* (representation of the auditory spectrum) is arranged along the cochlear duct. In other words, a characteristic resonance frequency can be determined at each point along the duct. The low-frequency resonators (low-pitched) are located near the apex, and the high-frequency resonators (high-pitched) are located at the base of the duct.



**Figure 1.11.** Tonotopy of the cochlear duct and frequency distribution

As well as being channeled through the air, sound vibrations are transmitted by means of another mechanism, known as *bone conduction*. Sound messages are directed to the inner ear via the bones in the skull. A much larger quantity of energy is required to produce a given stimulus via this path than is needed to propagate sound through the air. Indeed, the relative attenuation between the two paths has been measured as 30–60 dB, depending on the frequency.

### 1.2.2. Fletcher–Munson curves

The sensitivity of our ears is not linear with respect to the sound pressure. In other words, the perceived volume of a sound with a given intensity can seem higher or lower depending on the frequency of the signal. The *Fletcher–Munson curves* demonstrate this phenomenon. Below a certain frequency-dependent power threshold, sounds are imperceptible. This defines our threshold of hearing. The same is true for high-power sound messages, which become unbearable after a certain point, defining the threshold of pain. Fletcher established the curve relating the frequency on the *x*-axis to the power (sound pressure) on the *y*-axis. When listening *binaurally* (with both ears), the curves show which sounds create identical sensations. This work was standardized in 1961 to define so-called *loudness contours* (*isosonic curves*).

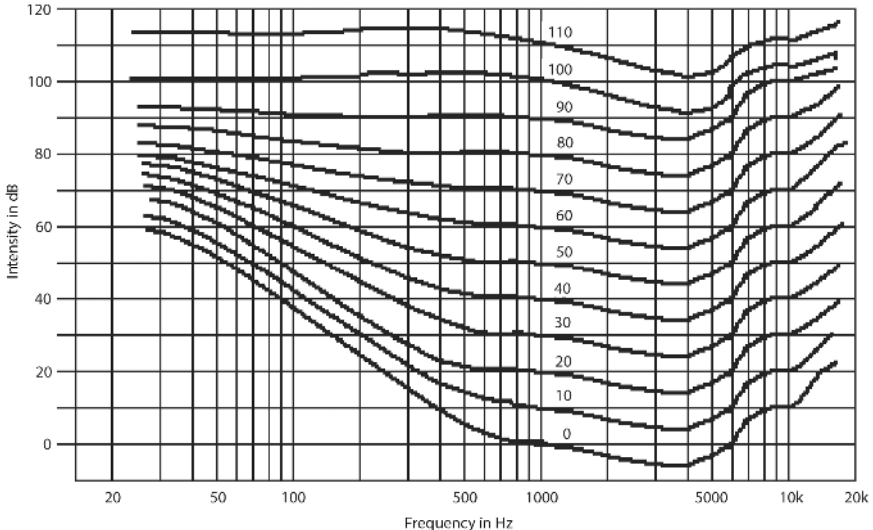


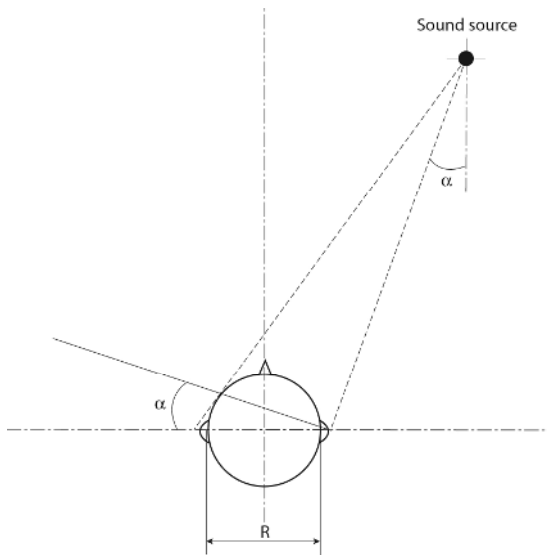
Figure 1.12. Fletcher–Munson curves (*isosonic curves*)

### 1.2.3. Auditory spatial awareness

Our ears are capable of pinpointing the source of a sound fairly accurately. This ability is based on several parameters.

In 1907, Lord Rayleigh<sup>5</sup> demonstrated the principles of *interaural level differences* and interaural time differences for the first time.

The *interaural time difference* (ITD) is a construction characterizing the time difference between the arrival of a sound wave at each of the two ears of a person. If the sound is coming from the front, this difference is zero.



**Figure 1.13.** Principle of sound localization by ITD

The ITD can be approximated with the equation:

$$\Delta_t = \frac{R(\alpha + \sin\alpha)}{c}$$

---

<sup>5</sup> John William Strutt Rayleigh, British physicist born at Langford Grove (Essex), 1842–1919. He was awarded the Nobel prize in physics in 1904 for the discovery of inert argon gas together with William Ramsay, and conducted extensive research into wave-related phenomena.

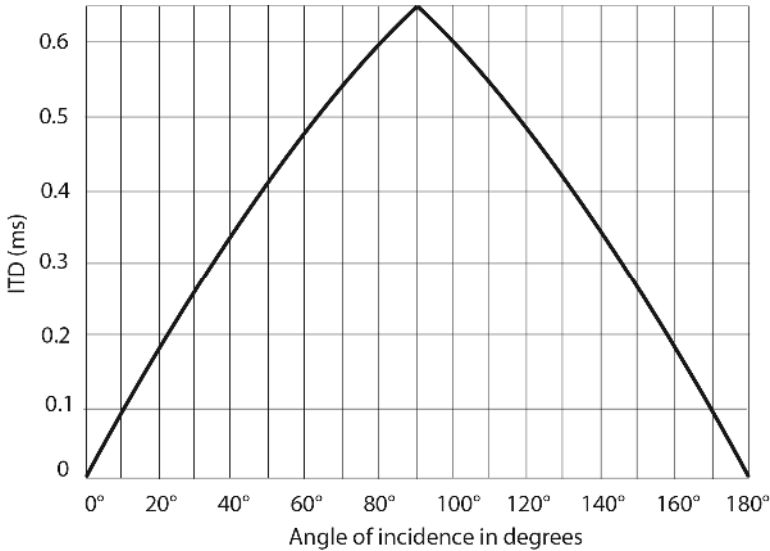
where:

- $\Delta_t$ : ITD in seconds;
- $R$ : radius of the head in meters (8.75 cm by default);
- $\alpha$ : angle of incidence in radians;
- $c$ : speed of sound (340 m/s).

For example, if a sound source is located at  $30^\circ$ , we find that:

$$\Delta_t = \frac{0.0875(0.523 + 0.5)}{340} = 2.63 \times 10^{-4} \text{s} = 263 \mu\text{s}$$

The sound will reach the person's ears with a time difference of 263  $\mu\text{s}$ . This difference can be thought of as a phase shift that is analyzed and interpreted by our brain in order to locate the position of the source. The maximum ITD is around 673  $\mu\text{s}$ .



**Figure 1.14.** *ITD calculation chart*

This phenomenon is particularly distinctive at low frequencies below 1,500 Hz. At higher frequencies, the interaural level difference (ILD), the interaural intensity difference (IID) or the interaural pressure difference (IPD) is used instead.

If a sound is emitted by a source that is closer to one ear than the other, there will be difference in the sound intensity or acoustic pressure. Our auditory system uses this difference to locate the sound source.

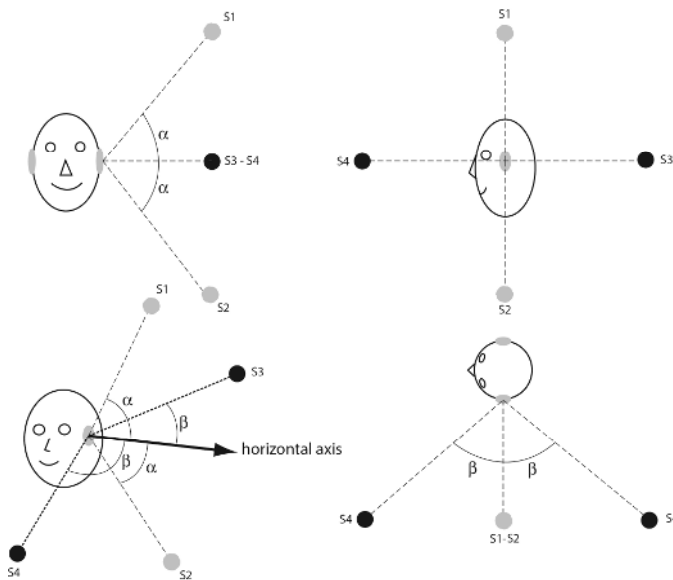
Gary S. Kendall and C.A. Puddie Rodgers proposed a simple equation to calculate the ILD:

$$\Delta_I = 1 + \left( \frac{f}{1,000} \right)^{0.8} \times \sin \alpha$$

where:

- $\Delta_I$ : ILD;
- $f$ : frequency in kHz;
- $\alpha$ : angle of incidence in radians.

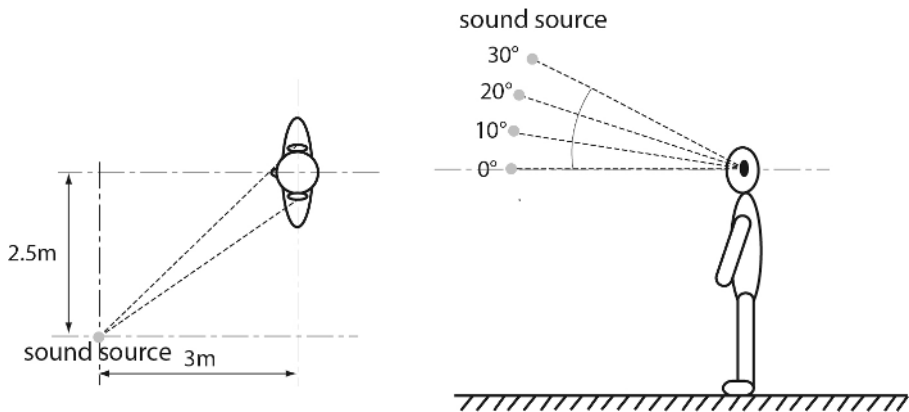
But the ITD and ILD alone are not the only factors that allow us to discriminate between the positions of different sources. In some cases, the ITD and the ILD may be identical, even though the sources are located at different positions, as shown in Figure 1.15.



**Figure 1.15.** Angular localization error in the horizontal and vertical directions. The sources  $S1$  and  $S2$  (vertical plane) have exactly the same ITD and ILD values as the sources  $S3$  and  $S4$  (horizontal plane)

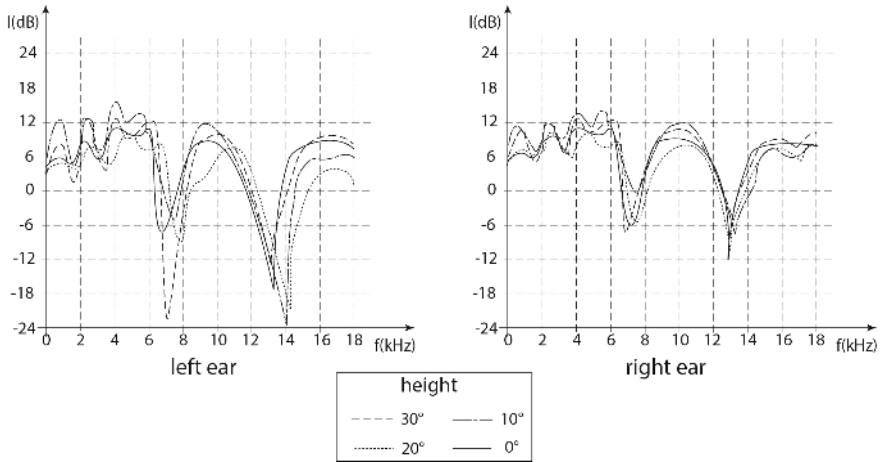
Another parameter is also used as a factor to locate the source of sounds based on diffraction caused by the morphology of the head. This eliminates the ambiguity created by the localization errors presented previously. Today, this factor is thought to be the most important factor, and is currently the subject of extensive research. The head-related transfer functions (*HRTF*) method represents the result of this work.

To understand how HRTF localization works, suppose that a source is placed directly to the right of an individual. His or her right eardrum will receive the sound message along a straight path. However, the sound waves that reach the left ear will need to follow a much more complex path around the head before ultimately hitting the left eardrum after multiple reflections and diffractions within the auricle and the ear canal.



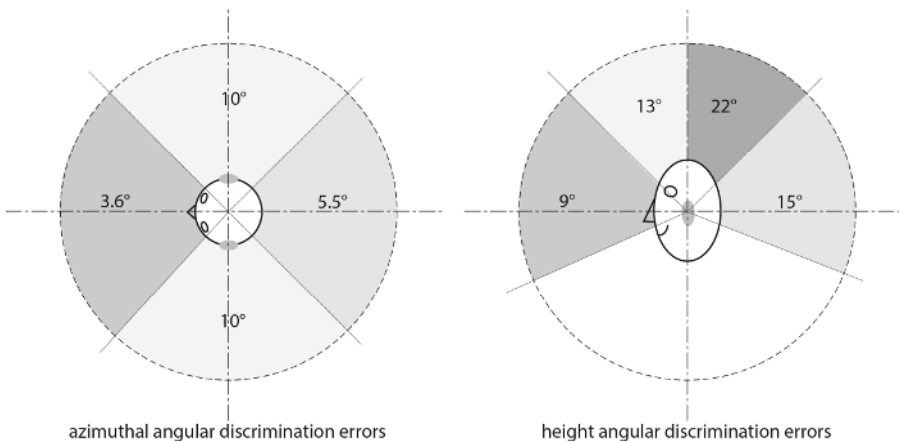
**Figure 1.16.** *Implementation of HTRF. The source is placed in front of the observer, and its elevation is varied from 0 to 30° relative to the horizontal plane through the observer's ears*

As it travels, the timbre of the sound wave, and therefore its spectrum, is modified. These modifications depend on the location of the source within three-dimensional space and are interpreted by the brain in order to determine this location. Note that HRTF localization is capable of determining vertical position, unlike the ITD and ILD.



**Figure 1.17.** *The HRTF can distinguish between different heights, unlike the ITD and ILD methods. The curves change as a function of the height of the source relative to the ears of the observer*

The angular localization errors<sup>6</sup> are shown in Figure 1.18, determined by all of the above methods (ITD, ILD and HRTF).



**Figure 1.18.** *Values of the angular perception errors in the horizontal and vertical directions*

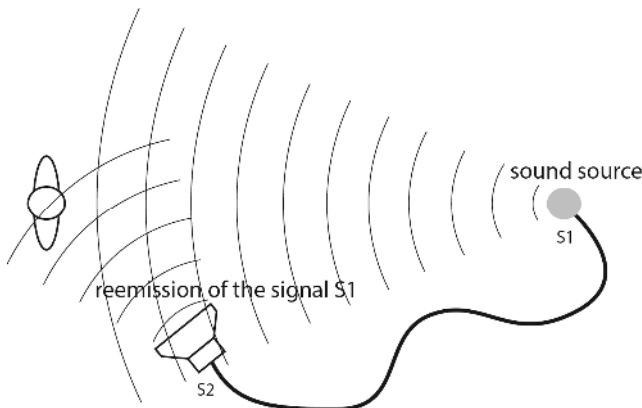
<sup>6</sup> Based on the measurements and experiments performed by Blauert.

It should be noted that the localization functions studied previously (ITD, ILD and HRTF) are based on binaural hearing (with both ears), but localization is also possible with monaural hearing. The auricle of the outer ear filters sound by means of reflections that depend on the angle of incidence of the source. These reflections introduce delays and timbre deformations, which can be used to determine the origin of the emitted sound.

Certain effects can impede or enhance the process of locating sounds in space. Two effects are particularly significant: the “cocktail party” effect, and the precedence effect, also known as the *Haas effect*.

The “cocktail party” effect occurs in noisy environments. We are capable of locating the source of a sound if we know where it is.

The Haas effect<sup>7</sup> reveals how we perceive reflected sounds compared to direct sounds. Our ears cannot distinguish between direct sounds and reflected (or reemitted) sounds if they are separated by less than 50 ms, even if the reflected source is louder<sup>8</sup> than the direct source by several dB. If the separation is longer, our ears no longer perceive the sources together, but instead as an echo. This phenomenon implies that our auditory system interprets the direction of an acoustic phenomenon as the direction of the first source that it perceives. This effect is also known as the law of the first wavefront.



**Figure 1.19.** Demonstration of the precedence effect or Haas effect. If the difference between the sources S1 and S2 is less than 50 ms, the observer cannot perceive a difference between them (no echo)

<sup>7</sup> According to a study conducted by Helmut Haas in 1951.

<sup>8</sup> Depending on the separation (0 to 50 ms), this value ranges up to a maximum of 10 dB.

### 1.3. The typology of sounds

Like everything that surrounds us, sounds can be organized into a typology based on characteristics distinguishable by physics or hearing.

#### 1.3.1. Sounds and periods

Pure sounds and constant frequencies rarely exist in our natural environment, except in certain circumstances. The sounds around us are typically complex. We can define a typology of sounds to distinguish them. The first thing to consider is the property of periodicity.

A sound is said to be *periodic* if its frequency does not change over time.

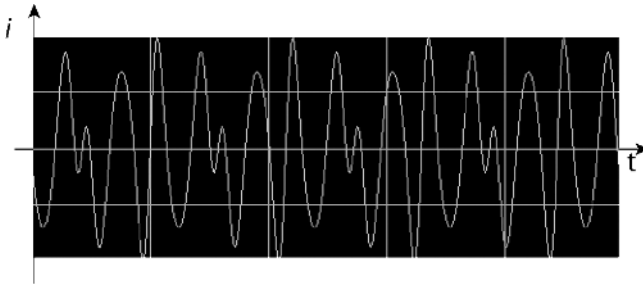


Figure 1.20. Representation of a periodic sound

A sound is said to be *aperiodic* if it is characterized by a large number of changes in frequency and amplitude over time. Most of the noises in our environment are aperiodic.

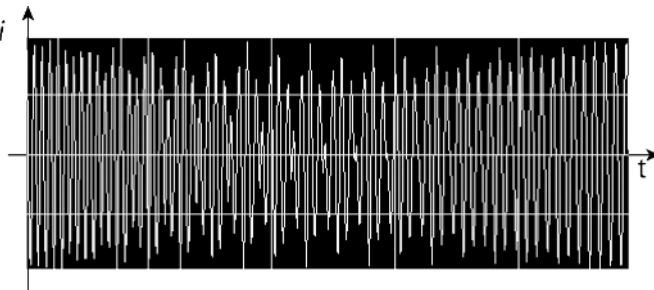
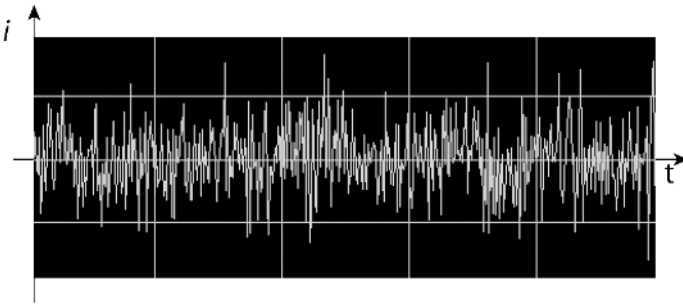


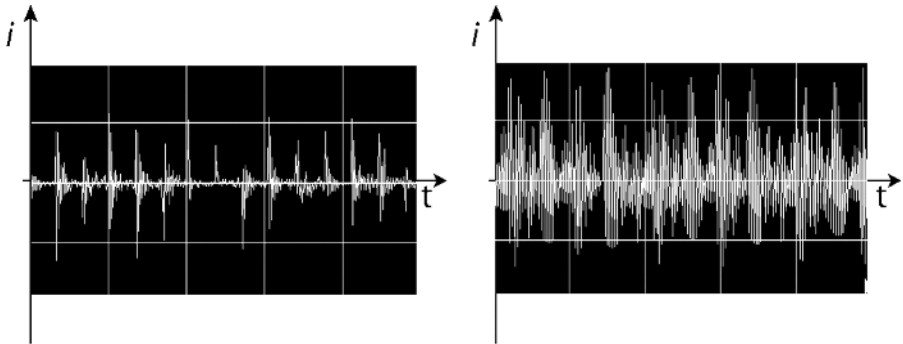
Figure 1.21. Representation of an aperiodic sound

*White noise* is an extreme example of an aperiodic sound. It uses the entire spectrum of audible frequencies.



**Figure 1.22.** Representation of white noise

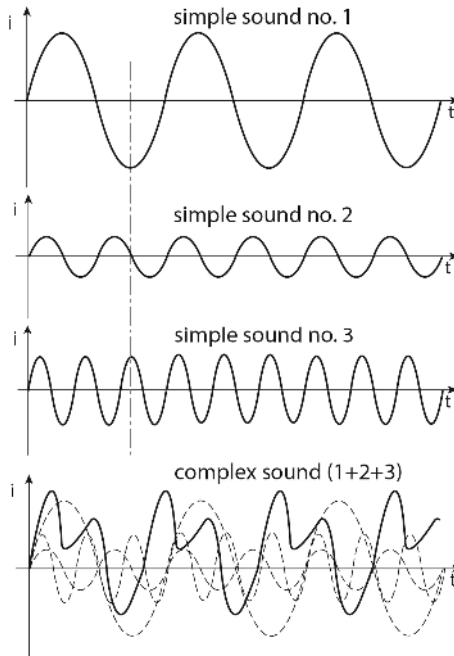
A sound with extremely short duration is called an *impulse*, whereas longer sounds are described as *continuous*.



**Figure 1.23.** Representations of impulses and a continuous sound

### 1.3.2. Simple sounds and complex sounds

Sounds with sinusoidal waveforms are called *simple sounds*. All other sounds are *complex sounds*. Complex sounds are composed of two or more sinusoidal waves. Thus, complex sounds are combinations of multiple simple sounds.



**Figure 1.24.** *Combination of multiple simple sounds*

To determine the simple waves that compose a *complex sound*, we can use a mathematical tool called a *Fourier transform*<sup>9</sup>.

Fourier transforms allow us to decompose complex sounds into a series of simple sounds (sinusoidal waves). Complex sounds in the real world can contain dozens of sinusoidal components.

The amplitude of a complex sound at any given moment in time is given by the sum of the amplitudes of the simple sounds from which it is composed.

The frequency of a complex sound is equal to the lowest of the frequencies of the simple sounds from which it is made up of.

The frequency of a complex sound is called the *fundamental frequency* ( $F_0$ ). If a sound is not periodic, it implicitly does not have a fundamental frequency. Instead, we think of it as a noise.

---

<sup>9</sup> Jean Baptiste Joseph Fourier, March 21 1768–May 16 1830, French mathematician and physicist.

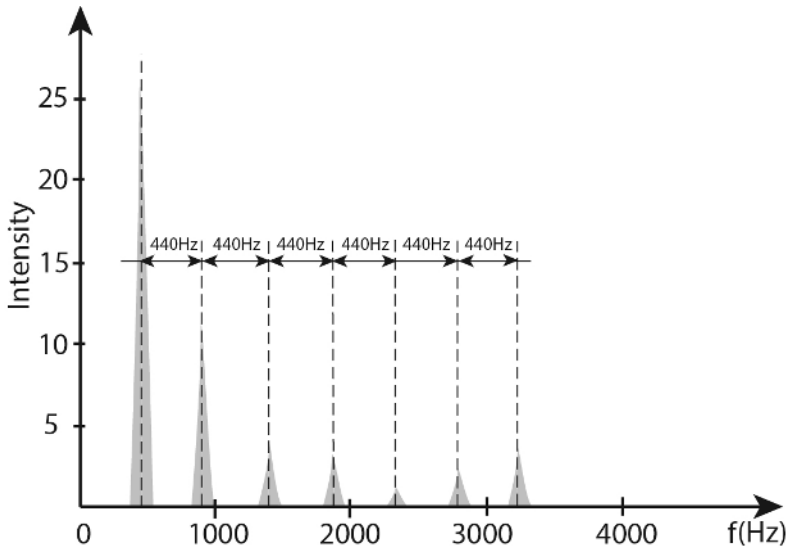
## 1.4. Spectral analysis

Spectral sound analysis combines several analytical techniques for determining the characteristics of an audio signal. The results of spectral analysis are often presented graphically.

### 1.4.1. The sound spectrum

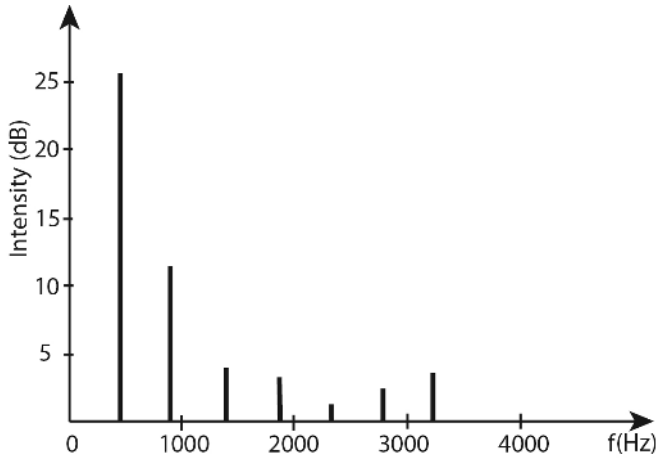
The spectrum of a sound is a representation of its amplitudes and frequencies. The representations that we used earlier always described each sound by a variation in amplitude along the vertical axis and a horizontal axis representing time.

The spectrum of a sound wave does not contain any information about time.



**Figure 1.25.** *Spectrum of a periodic complex sound*

The spectrum is often represented as a vertical bar graph (or line graph). Each bar shows the amplitude of a certain frequency. The lowest frequency is the fundamental frequency.



**Figure 1.26.** Line spectrum of the sound in Figure 1.22 (intensity in dB)

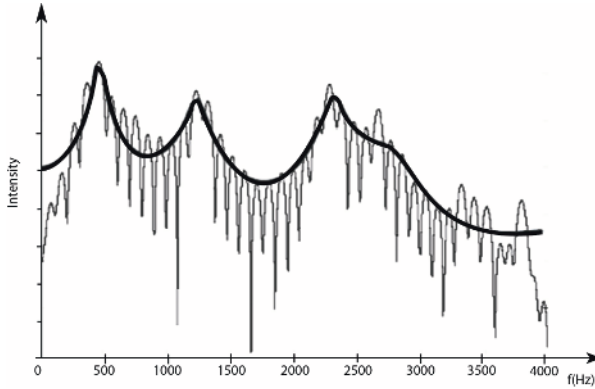
This type of graph is excellent for periodic sounds, since they contain a limited number of frequencies. We say that such sounds have a *discontinuous spectrum*, *line spectrum* or *comb spectrum*.

The other frequencies are called *harmonics* and are multiples of the fundamental frequency.

Designation	Symbol	Frequency
Fundamental frequency or first harmonic	$F_0$	440 Hz
Second harmonic	$F_1$	$2 \times F_0 = 2 \times 440 = 880$ Hz
Third harmonic	$F_2$	$3 \times F_0 = 3 \times 440 = 1,320$ Hz
Fourth harmonic	$F_3$	$4 \times F_0 = 4 \times 440 = 1,760$ Hz
Fifth harmonic	$F_4$	$5 \times F_0 = 5 \times 440 = 2,200$ Hz
Sixth harmonic	$F_5$	$6 \times F_0 = 6 \times 440 = 2,640$ Hz
Seventh harmonic	$F_6$	$7 \times F_0 = 7 \times 440 = 3,080$ Hz

**Table 1.2.** The harmonics of (440 Hz)

Sometimes we need to visualize the frequency distribution of the spectrum, in which case we use a graph showing the *spectral envelope* of the sound instead of a bar graph.



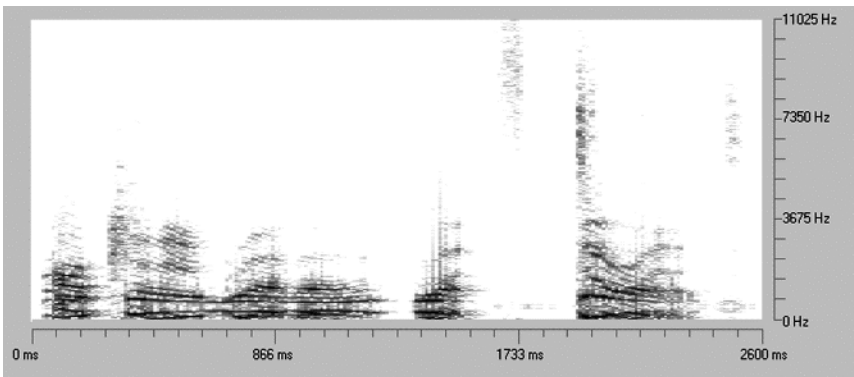
**Figure 1.27.** Spectral envelope of a complex aperiodic sound

This type of graph is excellent for aperiodic sounds with many frequencies. Sounds like this are said to have a *continuous spectrum*.

#### 1.4.2. Sonogram and spectrogram

One of the disadvantages of the spectral representation, as we saw earlier, is that it does not include any information about time. However, when the signal is *pseudo-periodic* (a sequence of periodic signals over time: musical phrase, speech, etc.), it can be useful to take time into account.

Accordingly, scientists developed new types of graph with three parameters (frequency, intensity and time), known as *sonograms* or *spectrograms*.



**Figure 1.28.** Sonogram of a sound sequence

A sonogram is a 2D graph where the sound intensity is defined by a scale of colors or shades of gray. A spectrogram is usually a 3D graph with time and frequency on the  $x$ - and  $z$ -axes, and the sound intensity on the  $y$ -axis.

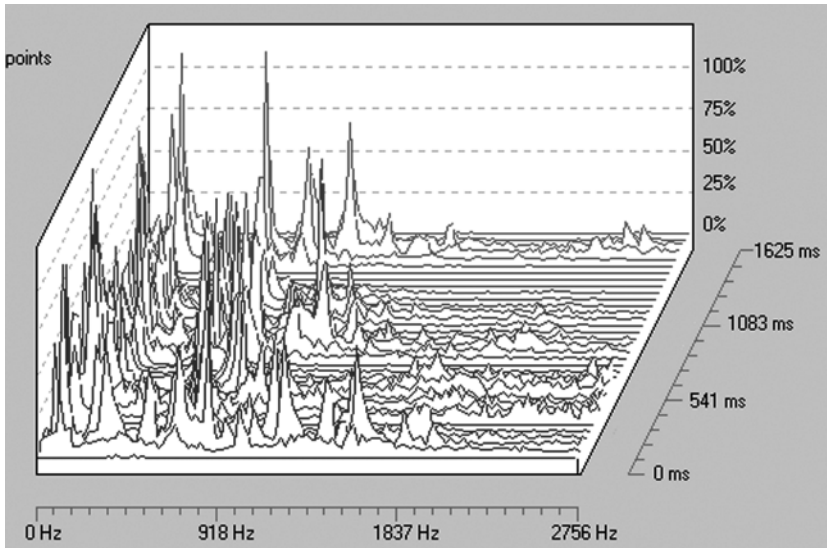


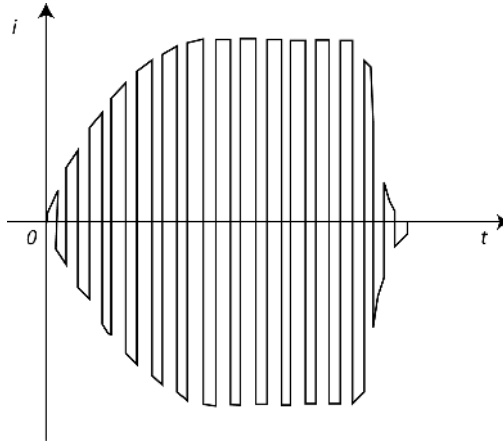
Figure 1.29. Spectrogram of a sound sequence

## 1.5. Timbre

The notion of timbre is too difficult to describe in terms of the fundamental frequency, the harmonics or the loudness (sound intensity). Timbre is a complex *psychoacoustic* concept. Sounds can be strident, dull, dry, warm and many more things.

### 1.5.1. Transient phenomena

Sound perception exists within a wider context. Each sound begins, stabilizes and then fades. Each single moment in the act of hearing a sound is a *transient phenomenon*. There are several types: *attack transients* and *release transients*, which describe the beginning and the end of sound phenomena. We can add other parameters to this list, which, like *vibrato* and *tremolo*, can either occur in a single phase (often when the sound has stabilized), or in every phase.



**Figure 1.30.** *The transients of the sound of a pipe organ*

The attack transient gives the sound its unique signature. For example, the attack transient is what allows us to distinguish between the sound of a clarinet and the sound of a flute. Its duration varies, typically ranging from 1 to 100 ms. It is thought that the human ear needs 40–50 ms to distinguish and recognize a sound. This type of transient is often complex. As well as its duration, we need to consider its *slope*, its instability, its spectral composition, the number of components and the order in which they occur, as well as many other factors.

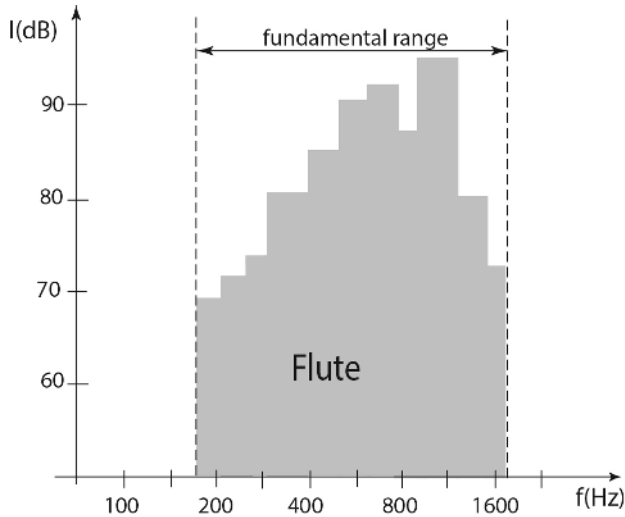
The stabilization phase (or sustain) is often steady, although it can be influenced by the environment or the musician's technique.

The release transients often depend on environmental parameters (echo, reverb, damping, etc.) and technique-related parameters for musical instruments. It can range from very short to very long (sound accessories or natural or artificial echo phenomena), from 1 to 5,000 ms.

Together, the transient phenomena: attack, sustain and release are often called the *envelope* of the sound.

### 1.5.2. Range

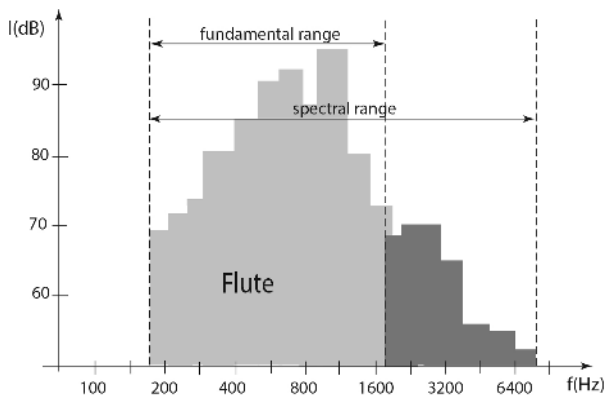
Each sound source, especially musical instruments and voices, can emit sound over a certain range of frequencies. This range is also known as the *tessitura*. However, we must be careful to distinguish between two concepts: the *fundamental range* and the *spectral range*



**Figure 1.31.** *Fundamental range of a flute*

The fundamental range (see Figure 1.31) only includes fundamental frequencies and harmonics. In an ideal setting, this is equivalent to the spectral range. However, reality often diverges from theory, which often attempts to describe the range of musical instruments simplistically as a set of fundamentals plus harmonics, with higher ranking harmonics having decreasing intensity.

The spectral range (see Figure 1.32) is the full spectrum of sounds that the source can produce.



**Figure 1.32.** *Fundamental range of a flute compared to its spectral range*

### 1.5.3. Mass of musical objects<sup>10</sup>

This notion of mass was introduced by Pierre Schaeffer<sup>11</sup>, who discussed it in his “Treatise on Musical Objects”, which remains an important reference to this day. On the subject of concrete sounds, i.e. sounds that are not affiliated with any musical instrument that is culturally recognizable to a given observer, he writes:

*But when we consider an arbitrary concrete sound (for example, produced by a membrane, a metal sheet, a rod...), we see that, unlike traditional sounds that have a clearly identifiable pitch, it has a certain mass located somewhere within its range, approximately characterizable by the intervals that it occupies, which are relatively easy to decipher. For example, it might contain several sounds with slowly changing pitch, which are dominated or surrounded by an aggregate of partials that are also gradually changing, the entirety of which can be approximately localized within a certain pitch interval. Our ears can quickly identify the most salient features and components, with some practice; these sounds then become as familiar to us as traditional harmonic sounds: they have a characteristic mass.*

His comments reveal the extent of the complexity involved in attempts to define the timbre. The concept of timbre extends far beyond the physics of the phenomenon into the realm of psychoacoustics. Schaeffer continues:

*At this point, we should stop to note that, since musicians care about things like: a note with good timbre, good or poor timbre, etc., they are distinguishing between two separate notions of timbre: one that relates to the instrument, indicating the provenance of the sound by the simple act of hearing, and another relating to each of the objects created by the instrument, involving an appreciation of the musical*

---

10 The concept of object was particularly important to Pierre Schaeffer: “A sound object refers to the studied signal, viewed within the context of perception. Instead of hearing events via sounds, we hear sounds as events.”

11 Pierre Schaeffer, 1910–1995, founder of the research department of the ORTF (Office de Radiodiffusion-Télévision Française), researcher in the fields of audiovisual communication and music. The inventor of musique concrète (concrete music), he created the GRM (Groupe de Recherche Musicale) in 1951, as well as being a composer (Variations sur une flûte mexicaine, Suite pour 14 instruments, Symphonie pour un homme seule, Toute la Lyre, Orphée 51, Masquage, and others), and author of the monumental and prophetic work “Traité des Objets Musicaux” (“Treatise on Musical Objects”).

*effects contained in the objects themselves, which are desirable in the context of musical interpretation and musical activities. We can go even further by speaking of the timbre of one single component of the object: the timbre of the attack, which is distinct from its stiffness.*

*However, given this definition, the timbre of an object is nothing other than the form and substance of its sound, its complete specification among the range of sounds that a given instrument can create, up to any admissible artistic variation. Associating the concept of timbre with the object therefore cannot help us describe the object itself any further, since it simply postpones the analysis of the subtleties of our qualified perceptions of the sound.*

#### 1.5.4. Classification of sounds

Based on the observations relating to the concept of timbre presented above, we can define a classification of different types of sound message.

Category	Composition	Spectrum	Example
Pure sound	Tonal sound without harmonics	Filiform	Sinusoidal sound (BC generator, synth VCO, etc.)
Tonal sound	Sound with an identifiable pitch	A band	Note played by an instrument (piano, harpsichord, violin, etc.)
Tonal group	Group of multiple tonal sounds	Multiple bands	Chord played by an instrument (piano, harpsichord, organ chord, etc.)
Nodal group	Aggregate sound without an identifiable pitch	Multiple bands	Set of percussions (multiple cymbals together)
Nodal sound	Set of multiple nodal groups	A band	Percussive sound (cymbal)
White noise	Group containing every pitch	Full spectrum	White noise generator
Complex sound	Mixed group containing tonal sounds, tonal groups, nodal sounds, nodal groups.	Complex shape	Natural sound (bell, gong, metal sheet, etc.)

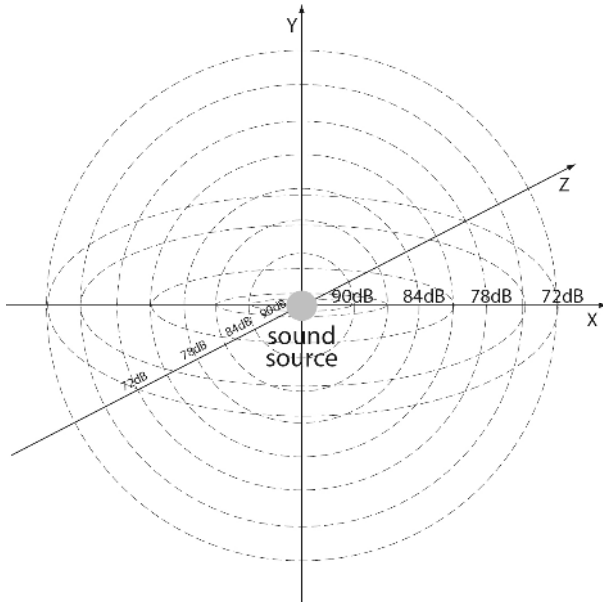
**Table 1.3.** Classification of sounds

## 1.6. Sound propagation

Sound waves propagate through their surrounding media by means of specific phenomena that result in specific behaviors. We will study the most important principles.

### 1.6.1. Dispersion

A sound wave emitted by a point source disperses as a set of concentric spheres.

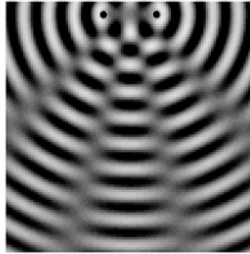


**Figure 1.33.** Spherical dispersion of a wave from a point source. The sound pressure level decreases by 6 dB whenever the distance doubles

Sound propagates through gaseous media like air, which is the most common transporting medium, in the form of alternating compressed and dilated layers. The set of wavefronts vibrate in phase (with each other).

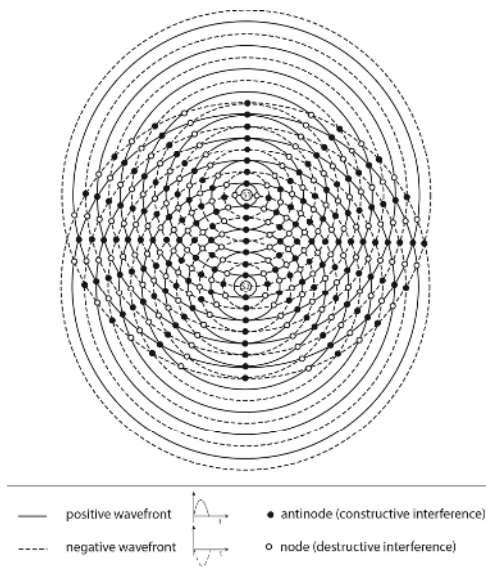
In order to describe certain phenomena, scientists introduced the abstract notion of plane waves, which do not actually exist in reality. Plane waves are just sections of spherical waves. When the source is sufficiently far away from the point of audition, spherical waves have a large radius of curvature, and so may be approximated by plane waves.

### 1.6.2. Interference



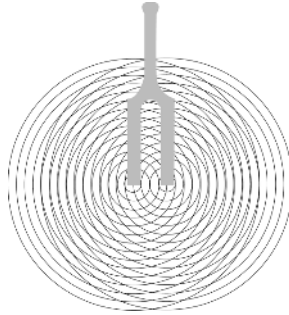
**Figure 1.34.** *Interference between two waves on the surface of a liquid*

When two sound waves meet, they overlap, and their intersection creates *constructive* or *destructive interference*.



**Figure 1.35.** *Interference between two identical waves (same frequency and amplitude). The sources S1 and S2 emit sound waves that overlap, creating interference (nodes and antinodes)*

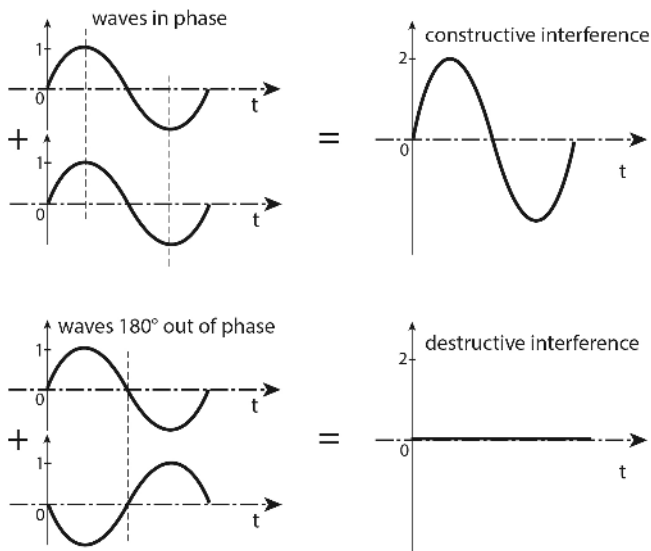
If you strike a tuning fork it and then rotate it near your ear, you will notice that it sounds louder or softer depending on the angle of rotation. This simple experiment demonstrates the creation of constructive and destructive interference.



**Figure 1.36.** *Interference created by a tuning fork*

When the sound waves emitted by each branch of the tuning fork are in phase, they add together to create constructive interference. But if they are perfectly out of phase, they generate destructive interference.

REMARK: There is an intermediate region between constructive interference (perfect sum of both signals) and destructive interference (perfect cancellation of both signals) in which the amplitude of the signal varies between its maximum and minimum values.



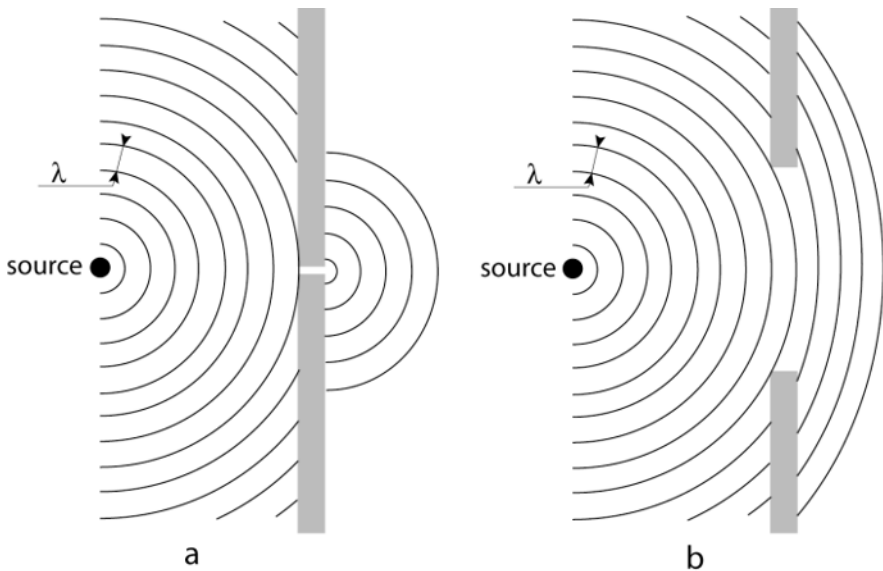
**Figure 1.37.** *Constructive and destructive interference*

### Constructive and destructive interference<sup>12</sup>

With two identical sound waves, the shape of the interference pattern is easy to determine. In the case of non-identical waves, it can be determined by the *principle of superposition*: when two sound waves occupy the same space at the same time, the total perturbation is simply the sum of each of the two individual perturbations at each point in space and time.

#### 1.6.3. Diffraction

When a sound wave encounters an obstacle, it goes around it. The edge of the obstacle becomes the center of a new wave, called a *secondary wave* or a *diffracted wave*.



**Figure 1.38.** Diffraction of a wave through an opening. The opening in a) is small relative to the wavelength  $\lambda$ , so a new point source is created. The opening in b) is larger than  $\lambda$ , so there is practically no diffraction

<sup>12</sup> We will return to the concepts of interference, phase and antiphase in section 4.1 when discussing equalization and filtering effects.

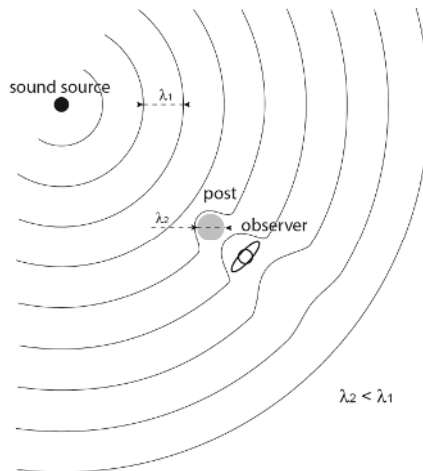
Diffraction is what allows us to hear sounds emitted by sources located behind obstacles.

Waves do not propagate in straight lines. Once a wave has moved past the obstacle, it again propagates in every direction. This phenomenon was demonstrated by the physicist C. Huygens<sup>13</sup>, who also gave us *Huygens' principle*.

“Each point on a wavefront may be viewed as the point source of a wave propagating in the same direction as the original wave. The next wave front may be obtained by summing all of these new waves”.

In simpler terms, a wave can be thought of as a sum of elementary waves vibrating in phase that are moving in the same direction as the original wave.

When the obstacle placed in front of the sound-emitting source is smaller than the wavelength, it does not exist from the perspective of the wave – the sound wave propagates around it as if it were “invisible”. The greater the wavelength-to-width ratio of the obstacle ( $\lambda/l$ ), the greater the effect of diffraction. When the ratio is large, an observer located behind the obstacle does not perceive any difference compared to when the obstacle is removed.



**Figure 1.39.** *Circumventing an obstacle smaller than the wavelength*

13 Christian Huygens, born at La Haye, Holland, April 14 1629–July 8 1695, mathematician, physicist and astronomer, author of “De ratiociniis in ludo alae” in 1657 and “Horologium oscillatorium” in 1673. He discovered Saturn’s ring, the rotation of Mars and the Orion nebula. He researched and published many papers on the wave theory of light, which enabled him to explain reflection and refraction, among other things.

If the width of the obstacle is large, part of the sound wave is reflected, and the edges of the obstacle become secondary sources of propagation.

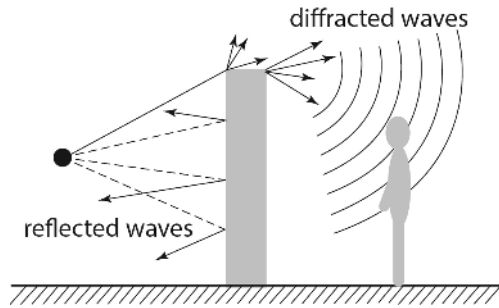


Figure 1.40. Emission of secondary waves at the edges of an obstacle

#### 1.6.4. Reflection

When a sound wave encounters an obstacle that it cannot circumvent because the width of the obstacle is greater than the wavelength of the wave, the wave is reflected. The *reflection* of sound waves follows the appropriately named *law of reflection* (Descartes' law<sup>14</sup>). This law states that the new direction of the sound wave after hitting a point on the surface is equal to the angle of the wave with the normal at this point at the moment of impact; in other words, the *angle of incidence* is equal to the *angle of reflection*.

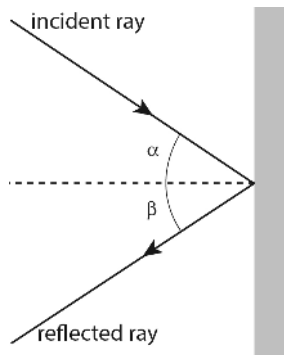
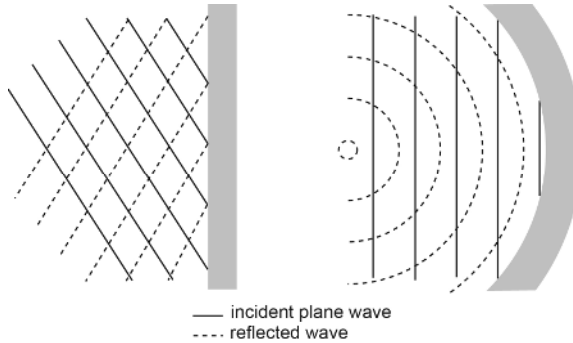


Figure 1.41. Angles of incidence and reflection of a sound wave ( $\alpha = \beta$ )

14 René Descartes, born in France, Touraine, 1596–1650, philosopher, physicist and mathematician, author of the famous “Discours de la méthode”, as well as his “Traité du Monde et de la Lumière”, “la Dioptrique”, “Météores” and so on.

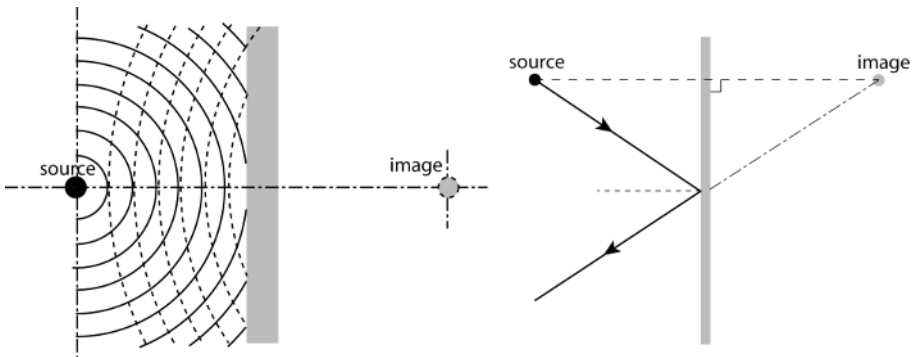
Note that the reflected waves can interfere with the incident waves to create constructive or destructive interference. Within a thin layer very close to the surface hit by the wave, the sound intensity is increased by the sum of the incident wave and the reflected wave. In this layer, the acoustic sound pressure is doubled, which increases the intensity by around 6 dB.



**Figure 1.42.** Reflection of a sound wave on a flat surface and a concave surface

If the obstacle reflecting the wave has a concave shape, a *focusing* phenomenon occurs. In the opposite case, when the obstacle has a convex shape, a *scattering* phenomenon occurs.

Wavefronts hitting a wall created reflected waves that act as if they were generated by an “image” of the sound source located the same distance on the other side of the wall.



**Figure 1.43.** Reflection of a sound wave on a wall and image of its source

When a sound wave strikes a hard surface, there is no phase change in the reflected wave. We say that the surface has an *acoustic impedance* greater than air. But when a sound wave encounters an obstacle with a lower impedance, there is a phase inversion (for example when a sound transmitted by a solid meets the air).

### 1.6.5. Reverberation (*reverb*)

This phenomenon will be discussed at great detail in Chapter 8, which presents time effects.

### 1.6.6. Absorption

When a sound wave hits an obstacle, it is reflected, but loses some of its energy. This energy is absorbed by the material. The ability of a material to capture or absorb sound energy is described by its *absorption coefficient*, often denoted as  $\alpha$ .

$$\alpha = 1 - |r|^2$$

where:

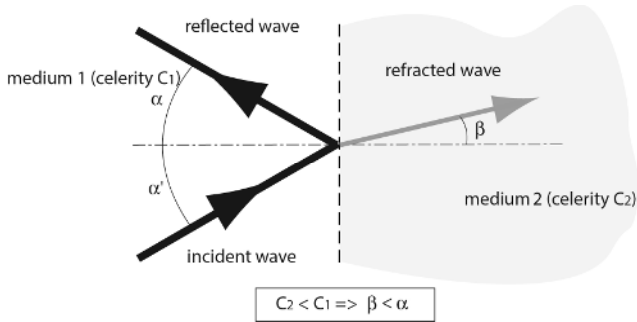
–  $\alpha$ : absorption coefficient (recall that  $\alpha = 1$ : perfectly absorbent,  $\alpha = 0$ : perfectly reflective);

–  $r$ : reflection factor.

Hard materials such as marble, cement and plaster have absorption coefficients ranging from 0.01 to 0.05, whereas porous or fibrous materials such as carpet, felt or glass wool have coefficients between 0.2 and 0.4 (see the table of absorption coefficients established by Sabine in section 8.1.1).

### 1.6.7. Refraction

When a sound wave passes from one environment to a different environment, it changes speed (*discontinuous celerity*) and direction. This generates both a reflected wave and a *refracted wave* (especially in the case of thin materials) with a lower energy than the original incident wave.



**Figure 1.44.** Refraction of a sound wave ( $\alpha = \alpha'$ )

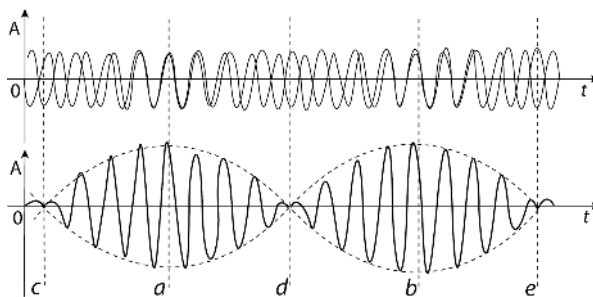
If the speed of propagation in the first medium is greater than the speed in the second medium, the angle of refraction is smaller than the angle of incidence or the angle of reflection.

### 1.6.8. The Doppler effect

This effect will be explained in section 5.4.2 of Chapter 5, which is dedicated to modulation effects, when we discuss the rotary effect.

### 1.6.9. Beats

The phenomenon known as a *beat* occurs when two sounds have similar frequencies. It is a direct consequence of the sensitivity of our auditory perception system. The alternating constructive and destructive interference of the sounds makes the volume of the sound signal seem to alternate between loud and quiet.



**Figure 1.45.** The beat phenomenon. At a) and b), the signals are in phase, at c), d) and e), they are completely out of phase (antiphase)

The frequency of the beat is equal to the absolute value of the difference of the two frequencies of each sound wave.

$$f_{beat} = |f_2 - f_1|$$

If the frequency of the beat is greater than 50 Hz, our brain is able to distinguish between the two sound sources. If not (<50 Hz), our brain detects a sound with an intermediate pitch (often called: the *subjective tone* or *combination tone*) with an intensity that appears to vary at the rate of the beat frequency.

## 1.7. Conclusion

Now that we have taken a brief tour around the notions of sound and acoustics, readers will have gained the foundations needed to understand the various sound effects that we shall discuss throughout the rest of the book.

We have not even nearly mentioned everything there is to know about sound. Interested readers can explore further using the web links and bibliography given at the end of this book, but you should now be able to continue reading without any difficulties, if you prefer.

