1

Sounds, Creation and Generation of Notes

To begin this book, which has the ambitious aim of serving as a passport to harmony in the musical domain, it is perfectly normal to offer a recap of a few elementary aspects, which are absolutely necessary for the workings of our auditory apparatus (ear + brain + education + civilization + etc.) which will, ultimately, be the adjudicator of all this work. Thus… 1; 2; 1, 2, 3, 4!

1.1. Physical and physiological notions of a sound

1.1.1. *Auditory apparatus*

In acoustic science, sound is a vibration propagating through gases, liquids and solids. For humans, generally, it is the vibration of a mass of air, driven by a tiny variation in air pressure, with varying rapidity, which, via the outer ear, vibrates the membrane of the hearer's eardrums and stimulates nerve endings situated in the inner ear (see Figure 1.1).

1.1.1.1. *Outer ear*

The auditory canal in the "outer" ear is in the shame of an acoustic horn, decreasing in diameter as we approach the bottom – i.e. the eardrum.

1.1.1.2. *Middle ear*

The middle ear contains the eardrum and three tiny bones, respectively called the hammer, anvil and the stirrup, which, together, make up the "ossicular chain". The hammer and the anvil form a fairly inflexible joint called the "incudomalleolar joint". The vibrations of masses of air in the auditory canal cause the eardrum to vibrate. These mechanical vibrations are then transmitted along the ossicular chain mentioned above, and then into the inner ear through the oval window.

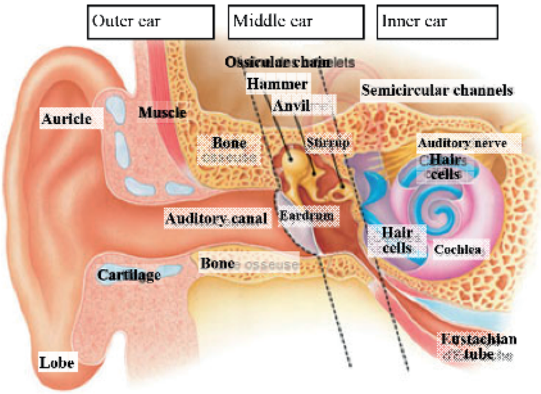

Figure 1.1. *Diagram of the human auditory apparatus (source: Wikipedia). For a color version of this figure, please see www.iste.co.uk/paret/musical.zip*

The mode of simplified propagation of vibrations in the inner ear is essentially as follows: the lines of the concentric zones of iso-amplitude of certain frequencies are parallel to the "handle" (shaft) of the hammer, with, for the membrane of the eardrum, zones of vibration with greater amplitude than the handle.

As the middle ear forms a cavity, overly high external pressure may perforate the eardrum. In order to ensure and re-establish a pressure balance on both sides of the eardrum (inner/outer), the middle ear is connected to the outside world (the nasal cavities) via the Eustachian tubes.

1.1.1.3. *Inner ear*

The inner ear contains not only the organ of hearing, but also the vestibule and the semicircular ducts, the organ of balance (not shown in the figure), responsible for perception of the head's angular position and its acceleration. Microscopic motions of the stirrup are transmitted to the "cochlea" via the oval window and the vestibule.

The cochlea is a hollow organ filled with a fluid called endolymph. It is lined with sensory hair cells (having microscopic hairs – cilia – which serve as sensors), which cannot regenerate once lost. They have tuft-like protruding structures: *stereocilia*. These cells are arranged all along a membrane (the *basilar membrane*), which divides the cochlea into two chambers. Together, the hair cells and the membranes to which they connect make up the "organ of Corti".

The basilar membrane and the hair cells are set in motion by the vibrations transmitted through the middle ear. Along the cochlea, each cell has a preferential frequency to which it responds, so that on receiving the information, the brain can differentiate the frequencies (the pitches) making up the different sounds. The hair cells nearest the base of the cochlea (oval window, nearest to the middle ear) tend to respond to high-pitched sounds, and those situated at its apex (final coil of the cochlea), on the other hand, respond to low-frequency sounds.

It is the hair cells which carry out mechanical- (i.e. pressure-based) electrical transduction of the original signal: they transform the motions of their cilia into nerve signals, sent via the auditory nerve. It is this signal which is interpreted by the brain as a sound whose tone height (pitch) corresponds to the group of cells stimulated.

Thus, we have briefly recapped the set (mechano-electric + brain) of our likes and dislikes, which it is important to please, and therefore stroke the hairs as much as possible in the direction of the grain!

Following this brief interlude, let us now return to simple physics.

1.1.2. *Physical concepts of a sound*

A sound is represented by a physical signal – a variation in pressure in the ear – whose characteristics are primarily represented by three parameters:

– the set of instantaneous frequencies making up the acoustic signal – in physical terms, the spectrum or spectral content (the frequency values are expressed in Hertz, representing a certain number of wave variations per second);

– the acoustic level, expressed in the form of pressure (Pascal), power (acoustic Watts or else transposed into decibels – dB) or in acoustic intensity (W/m^2);

– the respective evolutions of the amplitudes and spectra of the sound as a function of the time (temporal evolution) between the sound's appearance and its complete cessation (the conventional English terms are attack, decay, sustain and release time) – see Figure 1.2.

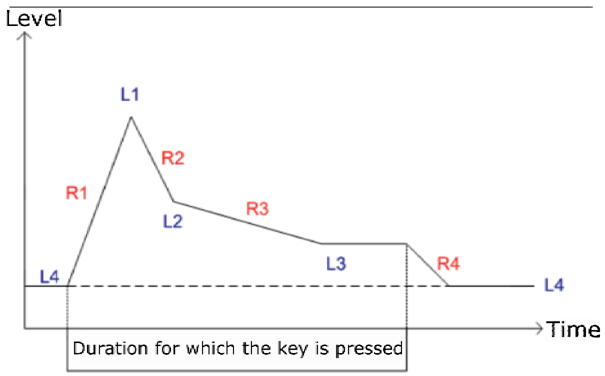

Figure 1.2. *Attack R1, decay R2 and R3, sustain L3 and release time R4. For a color version of this figure, please see www.iste.co.uk/paret/musical.zip*

The simplest and purest sound corresponds to a single-frequency sine wave, at constant amplitude, and sustained (see Figure 1.3). This is all very well, but hearers will very soon get bored of it!

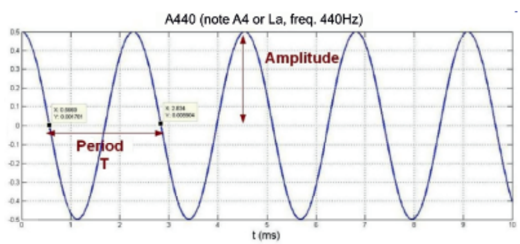

Figure 1.3. *Single-frequency, constant-amplitude, sustained sine wave*

1.1.3. *Further information on acoustics and acoustic physiology*

In order to help readers better comprehend and prevent unfortunate misunderstandings, here are few useful definitions.

1.1.3.1. *Energy (acoustic)*

By its very principle, a sound source diffuses acoustic energy E. Like any respectable form of energy, this is measured in Joules (J). The acoustic energy is linked to the movement of the air molecules propagating in the acoustic wave.

1.1.3.2. *Power (acoustic)*

If an acoustic source emits a sound with energy E (in Joules) over a time period Δt (in seconds), the acoustic power *Pow* of that source is the amount of energy emitted over that time period, and is defined as:

$$Power = E/\Delta t$$

The (acoustic) power is measured in Watts (W).

Examples of acoustic powers of a number of sources					
Normal voice	0.01 mW	=	0.00001 W	=	1.10^{-5} W
Loud voice	0.1 mW	=	0.0001 W	=	1.10^{-4} W
Shout	1 mW	=	0.001 W	=	1.10^{-3} W
Loudspeaker	1 W				
Airplane	1 kW	=	1000 W	=	1.10^{3} W

Table 1.1. *Acoustic power of several sources*

1.1.3.3. *Pressure (acoustic)*

The pressure *Pres* results from a force F (in Newton or kg per m/s²), applied to a surface S (in square meters, m²). Its value, therefore, is defined as:

$$Pres = F/S$$

The pressure is measured in Pascal (Pa) or Newton/meter² (N/m²).

In acoustics, we distinguish between two particular types of pressure values:

– atmospheric pressure “P_0” (also known as static pressure), which is the pressure exerted by the atmosphere (all molecules of air) on the Earth, on all humans and, of course, on their eardrums. Its value varies somewhat with changing weather conditions, but we can state that, on average, it is:

$$P_0 = 1.013\times10^5 \text{ Pa} \simeq 10^5 \text{ Pa}$$

– acoustic pressure “p” (also known as dynamic acoustic pressure). As a sound wave propagates, the air molecules which are set in motion cause slight local variations in the atmospheric pressure. This dynamic variation of pressure is what we called acoustic pressure “p”.

This acoustic pressure p exerts a new force on the eardrum. This causes it to vibrate, so it transmits the waves to the brain via the mechanisms of the middle- and inner ear, as described above.

Thus, at a given point in space, the total pressure is:

$$P_{tot} = P_0 + p$$

NOTE.– The atmospheric pressure P_0 is an ambient pressure, known as the "absolute" pressure, and therefore always positive, whilst the acoustic pressure p is a fluctuation around atmospheric pressure, and hence can either be positive (overpressure) or negative (pressure deficit).

1.1.3.4. *Intensity (acoustic)*

Acoustic intensity, or sound intensity, corresponds to a quantity of acoustic energy E (in Joules) which, over a period Δt (in seconds), traverses a surface area (be it real or virtual) S (in m^2). Thus, it is defined as:

$$I = (E / \Delta t) / S) = Pow / S$$

Thus, the acoustic intensity is measured in W/m^2.

If we suppose that the sound source radiates uniformly in all directions in space (i.e. it is a source said to be isotropic, or homogeneous), it will emit waves with spherical wavefronts. If the radius of the sphere is *r*, its surface area will be equal to $S = 4\pi r^2$ and the acoustic intensity received over the spherical whole of a wavefront would be equal to:

$$I = Pow / 4\pi r^2$$

with Pow being the acoustic power of the source.

Thus, in the case of an isotropic source with power *Pow*, the acoustic (sound) intensity decreases in inverse proportion to the square of the distance *r* from the source. In addition, the sound reception of the human ear to the amplitude of a sound is not directly proportional to its amplitude, but is proportional to its logarithm. Thus:

– doubling the power of the sound source is equivalent to increasing the sound level by 3 dB;

– quadrupling that power increases the level by 6 dB;

– multiplying the power by 10 increases the level by 10 dB;

– multiplying it by 100 adds 20 dB to the initial level.

1.1.3.5. *Acoustic propagation, pressure, velocity and impedance*

Without wishing to inundate the reader with complex mathematical formulae, for general culture, note that the equation of the displacement "a" of the molecules which constitute the medium in which the sound wave is circulating is in the form of a second-order partial differential equation with respect to the distance ($\partial^2 a/\partial x^2$) and the time ($\partial^2 a/\partial t^2$), known as *d'Alembert's formula*. It is written:

$$\partial^2 a/\partial x^2 - (1/c^2)\,(\partial^2 a/\partial t^2) = 0$$

"c" represents the celerity (velocity) of the wave and depends on the nature of the medium.

This equation can easily be solved in a number of simple cases – particularly in relation to steady-state waves along an infinitely long tube. A solution is a function of the two aforecited variables *a, t* of the type:

$$a(x,t) = a_0 \sin(\omega t - kx)\ldots$$

which is a classic equation to describe a propagation phenomenon.

Furthermore, it is also possible to apply this equation of motion $a(x,t)$ to the propagation, to the acoustic pressure $p(x,t)$ or to the rate of vibration $v(x,t)$ of the molecules in the medium.

In the case of a plane soundwave (as can be observed in a pipe or in the canal of an ear whose diameter is less than half the wavelength of the sound vibration) and for low amplitudes, the acoustic pressure $p(x,t)$ and the acoustic vibration rate $v(x,t)$ of the associated particle of the medium vary together, are in phase and are linearly linked by the relation:

$$p(x,t) = (\rho c)\, v(x,t)$$

$$p = (\rho c)\, v$$

where:

p: pressure in Pa;

v: velocity of vibration in m/s;

ρ: density of the medium in kg/m³;

c: velocity of the sound wave in the medium in m/s.

1.1.3.5.1. Acoustic impedance

Stemming directly from the above equation, for an acoustic wave, we define the acoustic impedance Z_{ac} of a medium as being the ratio between the acoustic pressure and the velocity of the associated particle of the medium:

$Z_{ac} = p/v$
The acoustic impedance is measured in Pa·s/m.

The acoustic impedance is expressed in Pa·s/m, (also called the "rayl" in honor of John William Strutt, 3rd Baron Rayleigh), and for a progressive acoustic plane wave, it is therefore equal to (see previous section):

$Z_{ac} = +/- \rho c$

The sign depends on the direction of propagation and on the choice of orientation of the axis of propagation of the sound wave. As a characteristic property of the medium, the product (ρc) representing the impedance often has greater acoustic importance than ρ or c taken in isolation. For this reason, the product (ρc) is also known as the medium's specific characteristic acoustic impedance.

1.1.3.6. *Relation between intensity (W/m²) and acoustic pressure (Pa)*

As stated above, a propagating wave gradually transports energy.

The (total) energy *e* of a mass *m* oscillating on either side of a position of equilibrium is, at each time, the sum of its kinetic energy and its potential energy. In the absence of friction, the energy *e* is time-independent:

– when the velocity v is zero, the kinetic energy, $\frac{1}{2} mv^2$, is zero and the maximum potential energy;

– conversely, when the velocity passes through its maximum, "v_0", the kinetic energy is maximal and the potential energy is zero.

At every moment, therefore, the total energy of the particle is:

$e = \frac{1}{2} (mv_0^2)$

with *e* in J, *m* in kg, and v_0 in m/s.

The volumetric energy density *E* (the energy contained in a unit volume) will therefore be:

$E = \frac{1}{2} (\rho v_0^2)$

where ρ is the mass per unit volume = density of the medium in kg/m³.

However, the energy contained in the wave propagates with a celerity c. As shown by Figure 1.4, we can deduce that the "total energy" E which, over the "unit time" Δt, crosses the "unit surface" S (called the surface power or acoustic intensity, I) is the energy contained in a volume whose base has a unit surface and whose height is equal to c.

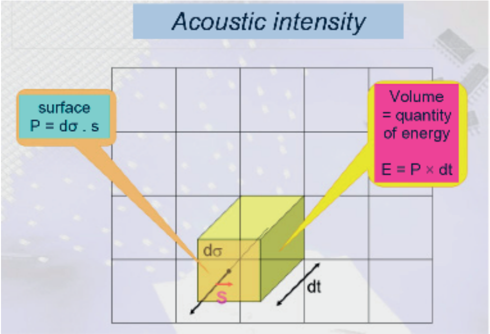

Figure 1.4. *Graphical representation of the acoustic intensity. For a color version of this figure, please see www.iste.co.uk/paret/musical.zip*

From this, by definition, it follows that:

$I = (E / \Delta t) / S) = Pow$ / S, acoustic intensity in W/m²

$I = cE = ½ (\rho c v_0^2)$

$I = (\rho c v_0^2) / 2$

Taking account of the following (see above):

$p = (\rho c)\ v_0$

"I" can be expressed in the following better-known form:

$I = p_0^2/2\rho c$

The total amount of sound energy J delivered to the tissues by a sonic irradiation over a time-period Δt is therefore:

$J = I\ \Delta t$

The pressure is a *force* per unit surface which is exerted on a *mass*: that of the neighboring element, and endows it with an *acceleration* – i.e. a variation in velocity.

In the conditions set out in the above paragraphs, we can show that at a certain point in space, the acoustic intensity I is linked to the acoustic pressure by the relation:

$$I = p^2 / \rho_0 c$$

in W/m^2, where:

– p: acoustic pressure in Pa;

– ρ_0: density of the medium (in air, $\rho_0 \simeq 1.2$ kg/m^3 in normal conditions of temperature, humidity and atmospheric pressure);

– c: celerity (velocity) of sound and c = 340 m/s in the same conditions;

Thus:

$$I \approx p^2 / 400$$

in W/m^2, with p in Pa.

NOTE.– In free space, the waves are not flat-fronted, but spherical, and the pressure and velocity do not vary exactly with one another. However, the more the radius of the sphere grows, the more the wave resembles a plane wave. The calculations with sine waves show that when the distance to the source is greater than or equal to the wavelength, we can assimilate the spherical wave to a plane wave.

NOTE.– This relation is valid only for direct sound coming from an acoustic source; not for sound reflected or echoed in a room.

1.1.3.6.1. Hearing- and pain thresholds

Whilst this may seem a little out of our field of study, let us say a few words about these two thresholds; in a few paragraphs, their relevance will become apparent, with the phenomena of masking as they also relate to concepts in harmony.

Hearing threshold

The hearing threshold corresponds to the weakest sound (in terms of intensity) than an “average” ear is capable of perceiving. It corresponds to a mechanical vibration of the eardrum of around 0.3×10^{-10} m, which is minuscule.

At 1000 Hz, which is the frequency value commonly accepted for that referential hearing threshold, the corresponding acoustic pressure is:

$$p\_ref = 2\times10^{-5}\ \text{Pa} = 20\mu\ \text{Pa}$$

In the knowledge that $I = p^2/\rho_0 c$ (in W/m^2), the hearing threshold corresponds to a sound intensity of:

$$I\_ref = 1\times10^{-12}\ W/m^2 = 1\ pW/m^2,\ \text{i.e. } 0dB(A)$$

Pain threshold

The value of the pain threshold corresponds to the acoustic pressure with maximum intensity that the “average” ear can withstand without being damaged. The commonly accepted value is:

$$p\_pain = 20\ \text{Pa}$$

Knowing that $I = p^2/\rho_0 c$ in W/m^2, the pain threshold corresponds to a sound intensity of:

$$I\_pain = 1\ W/m^2,\ \text{which is } 120\ dB(A)$$

Thus, the ear has a dynamic (operational range) of around 120 dB (corresponding to acoustic intensities from 1 pW/m^2 to 1 W/m^2).

1.1.3.7. *Fletcher–Munson curve*

Now that all these units are clearly defined, we can go beyond the physics and the beautiful mathematics and examine the hearing of a sustained sound. In this case, the human ear has specific characteristics that are summarized by the Fletcher–Munson curve / diagram (see Figure 1.5), which constitutes the auditory apparatus’ response to the stimulus of frequency as a function of the power emitted.

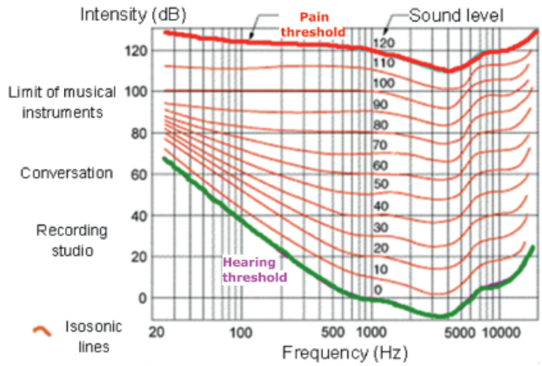

Figure 1.5. *Fletcher–Munson diagram / curve. For a color version of this figure, please see www.iste.co.uk/paret/musical.zip*

This figure shows the curves of equal sound sensations of a normal, functioning human ear as a function of the frequency and indicates that the range of normal hearing lies between the hearing threshold taken as a reference (i.e. 0 dB at 1000 Hz, which corresponds to a sound intensity of 10^{-12} W/m²) and the pain threshold (around 120 decibels). Furthermore, this hearing range is limited to around 30 Hz at low frequencies and around 15,000 Hz for high frequencies. In addition, it shows that the sensitivity of the ear is greatest in a range between around 1 and 5 kHz. It is well known that these limits vary from one subject to another, and even in the same individual, as a function of his/her age or diseases and accidents.

In the frequency range audible by the ear (typically roughly 20-20,000 Hz, but this differs from individual to individual), any audible “frequency/amplitude” couple on that diagram (but for the time being, the influence of amplitude is not considered to be critical) can be assimilated, simplistically, for now, to a “note”. On principle, therefore, it is/will be easy to generate “notes” by creating a succession of frequencies – arbitrarily or at random – or indeed by defining a particular or specific mode of creation of the succession of those frequencies. Nonetheless, these “notes” may have physical relations (physical values) with one another, particularly in terms of the respective values of their frequencies. We shall examine this in detail in subsequent chapters.

1.1.4. *Idea of minimum audible gap/interval between two frequencies*

Once again, the intrinsic qualities (mechanical – eardrums, physiological – brain, etc.) of the ear and of the whole auditory apparatus are involved.

1.1.4.1. *Acoustic frequency resolution*

Indeed, before going further, it is necessary and important to quantify – physically and physiologically (measure/define) – what might be the frequency resolution of an "average" ear; in other words, its auditory separation ability – i.e. which is the smallest frequency gap with which it is capable is distinguishing between two nearby frequencies.

We can show that the acoustic resolution of the human auditory apparatus depends on numerous factors. Among the main ones, we can cite:

– the absolute and relative acoustic levels at which the two frequencies involved are perceived;

– the presence or absence of neighboring frequencies of large amplitudes (see below):

- either simultaneous (frequency masking effect);

- or slightly shifted in time (temporal masking effect);

– the type of modeling of the psycho-acoustic apparatus;

– phenomena and definitions of consonance and dissonance, and of harshness, which we shall discuss in detail in Chapter 5;

– etc.

In conclusion, let us simply state that there comes a time when the ear is no longer able to distinguish between two close frequencies.

To clarify, let us also state that an average ear is able to quite easily distinguish between over thirty frequencies with the range of an octave (from f1 to 2×f1 – see below). Therefore, should we choose to use it, this minimum gap could represent the smallest frequency gap between two successive notes.

NOTE.– Physicists/acoustics experts, for their part, have long been able to divide the octave into 301 values, each one representing a "savart" (see Chapters 4 and 5). Thus, human frequency resolution corresponds to around 10 savarts, or slightly less. Certain animals perform far better than we do in this area.

Let us go into a little more detail about all this.

As previously noted, it has long been known that the ear has maximum sensitivity in the frequency range between around 1 and 5 kHz. In addition, the sensitivity curves in the Fletcher–Munson diagrams represent the audibility threshold of a signal “A” as a function of its frequency *in the absence of an “interference” signal* if that interference is audible because it is above the perception threshold (see Figure 1.5, again). The question which then arises is: what happens when two signals are separated by a brief time lag, or are present simultaneously.

1.1.4.1.1. Static (in sustained sounds, of constant amplitudes but separated by a time lag)

In this case, the auditory apparatus (including the memory) must determine whether, at a given amplitude, one frequency is different or equal to another by successively switching between the two and listening to them one after the other. This gives an initial idea of the concepts of consonance and dissonance and of static frequency resolution of the auditory apparatus.

1.1.4.1.2. Dynamic: masking (in simultaneous sustained sounds of constant amplitudes)

Here, we enter into the zone of masking.

The concept of sound masking is closely linked to: the perception of a sound’s acoustic intensity; that of its frequencies (pitches), which we looked at in the sector on the physiology of the cochlea (notably with the outer hair cells); and to an auditory “channel” (i.e. a fiber in the auditory nerve) which responds only to a stimulus situated in a precise frequency range. This is what is known as the ear’s frequency selectivity.

“Frequency masking” or “simultaneous masking”

The phenomenon of frequency masking occurs when a “tested” signal and the “mask” signal are presented at exactly the same time. The louder sound masks (prevents us from hearing) another, quieter one – particularly if the values of their frequencies are close together. In this case, in the presence of multiple signals, the audibility threshold curve represented in this diagram is affected locally. For example, in the simple case of the simultaneous presence of two signals of neighboring frequencies, the presence of the louder signal raises the level of the audibility threshold in its vicinity, making the ear less sensitive to very nearby frequencies – hence the term “frequency masking”.

Figure 1.6 illustrates this scenario, where signal A, which was previously audible, is now masked by the similar signal B, which is more powerful than A.

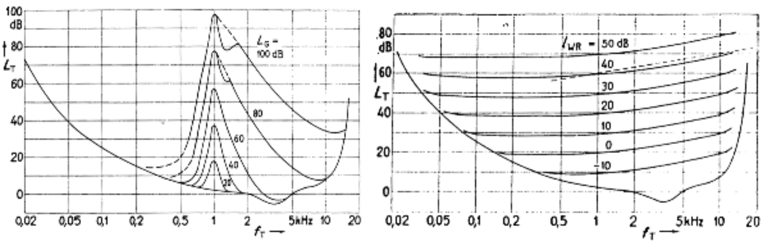

Figure 1.6. *Auditory impact of a loud signal B on a quiet similar signal A*

This frequency selectivity is measured using objective tests and gives the neural tuning curves. These tuning curves show the detail of these masking phenomena.

"Temporal masking" or "sequential masking"

There is also another type of masking, known as "temporal masking". This occurs when a high-amplitude sound masks weaker sounds that come immediately before or after it. Such masking takes place, in certain conditions, between two sounds which are not simultaneous, but instead are separated by a brief interval (e.g. the sound of a triangle immediately following the sound of a kettledrum). We then speak of temporal- or sequential masking, as opposed to the common case of frequency- or simultaneous masking discussed above.

Sequential masking manifests in short-term temporal interactions (spanning a few tens of a millisecond) between two stimuli/excitations. It is said to be:

– *proactive*, when the mask is presented before the test, meaning that the "mask" signal precedes the tested sound, which is also masked. This is the most significant case. It demonstrates mechanisms of inhibiting the excitability of the cochlea by way of an excitation immediately prior;

– *retroactive*, when the mask is presented after the test, meaning that the mask signal follows the tested sound, which is also masked. This masking, which could be described as "anti-causal" in signal processing, can only be explained as the interference of the temporal integration of these two competing signals.

Masking- and excitation patterns

To return to the subject that concerns us, in the Fletcher–Munson diagram, the zone/area of the "frequency/amplitude" plane in which another sound will no longer be perceived in the presence of the "mask" is known as the masking curve or "masking pattern" for that type of mask. The masking pattern of a pure sound or a narrowband sound exhibits a steep slope on the deeper-pitched side, and a gentler slope on the higher-pitched side. In this zone, therefore, masking is greater, which can be expressed by the statements "lower sounds mask higher ones" (see Figure 1.7).

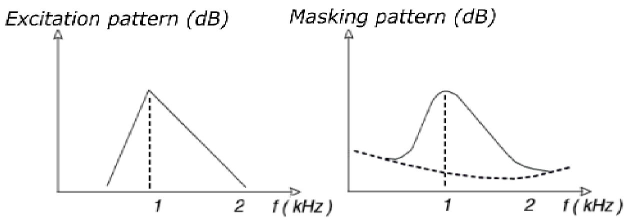

Figure 1.7. *Excitation pattern and masking pattern*

At the physiological level, the loud sound or "mask" produces a greater or lesser response in the various auditory channels near to its specific frequency or frequencies. The envelope of that response constitutes its "excitation pattern", which can be assimilated to the envelope of the vibrations of the basilar membrane. Its form can be explained (partly) by selectivity of the frequency of the deformations in the basilar membrane, reinforced by the active mechanisms in which the outer hair cells are involved. If it is strong enough, the excitation of the mask covers that produced by the quieter sound, which is therefore "masked". This interpretation implies that the masking pattern approximately coincides with the mask's excitation pattern. This hypothesis has been corroborated by tests and physiological observations.

To quantify these phenomena, one of the pertinent dependent variables is the SOA (Stimulus Onset Asynchrony), which is the sum of the time at which the test is presented and the duration of the time interval before the mask appears.

1.1.4.2. *Psycho-acoustic model of human hearing*

Within the group of experts – specialists in physiology, acoustics and physics – forming the devoted "Audio" team, the MPEG (*Moving Picture Experts Group*) in charge of writing the famous layer III of the MPEG 2 standard, well known as "MP3" sound encoding. After numerous experiments, the group chose a psycho-acoustic model of human hearing, which serves as the basis of the design of a so-called "perceptual" encoder. Such an encoder is characterized by a masking curve and a variable quantification that is a function of the ear's sensitivity in relation to the frequency instituting, and if two frequencies of different intensities are present at the same time, one may be perceived less than the other depending on whether or not their two frequencies are close together. This principle of modeling of our hearing is basically empirical, but fairly effective. In that percussive sounds (such as piano, triangle, drums, etc.) engender no perceptible pre-echo artifacts, it exploits the psycho-acoustic characteristics of temporal masking and "subband filtration", which consists of dividing the audio bandwidth (roughly 20-20,000 Hz) into 32 subbands of identical widths using an array of so-called "polyphase filters". This is the principle of "perceptual audio coding", which was chosen for compression into MP3 format to reduce the digital flowrate of information.

Following these highly scientific documentary points, we still need to make a number of additional comments before we can get back, in earnest, to what musicians expect from this book.

The audible band from 20 to 20,000 Hz, on principle, represents ten octaves, known as octave "0" to "9" (20 *40 80 160 320 640 1280 2560 5120 10240* 20480), which is often much too broad for a human being with a "typical" sense of hearing. In reality, humans typically only hear the eight main octaves indicated in italics here.

Additionally, in practice, the keyboard of an acoustic piano is usually made up of (81) 85 black and white keys (ranging from *la0* (or A0) to *la7*), and that of digital pianos has 88 keys (from *la0* to *do8*), which is slightly less than 8 octaves (see the example of the keyboard of a piano (Figure 1.8) and that of a conventional organ spans from *fa2* = 87 Hz… *la4* = 440 Hz… to *do9* = 8372 Hz).

NOTE.– As we shall see later on, the white keys correspond to the seven notes in the diatonic scale of *do* major) (C) and the black keys to the five remaining notes needed to make up a chromatic scale – 12 notes per octave in total.

If we maximize, considering a keyboard with 8 octaves of 12 notes per octave, this makes a total of 96 notes. We then take an array of 32 filters: this gives us a group of 3 notes per filter, so a filter bandwidth of one-and-a-half tones = 3 semitones. Thus, it is on the basis of a 3-semitone interval that frequency masking is used in signal analysis.

NOTE.– In MPEG 3, the analog signal (the sound) output by a subband filter is then sampled. By processing that signal, we are able to eliminate the subband signals below the threshold (not perceived by the hearer) in the psycho-acoustic model, and defines the precision of quantification needed for each of the subbands, so that the quantification noise remains below the audibility threshold in that subband. In addition, the zones where the ear is most sensitive can therefore be quantified with greater precision than the others.

1.1.5. *Why have we told this whole story, then?*

That is a good question! The answer is simple: why both adding in harmonic elements or ornaments into a piece of music when no-one will hear them because of consonance, frequency- and/or temporal masking, etc. Thus, it is important to always pay attention to these latent phenomena.

We have finished, for the time being, with the physiological aspect, so let us now look at how notes are generated.

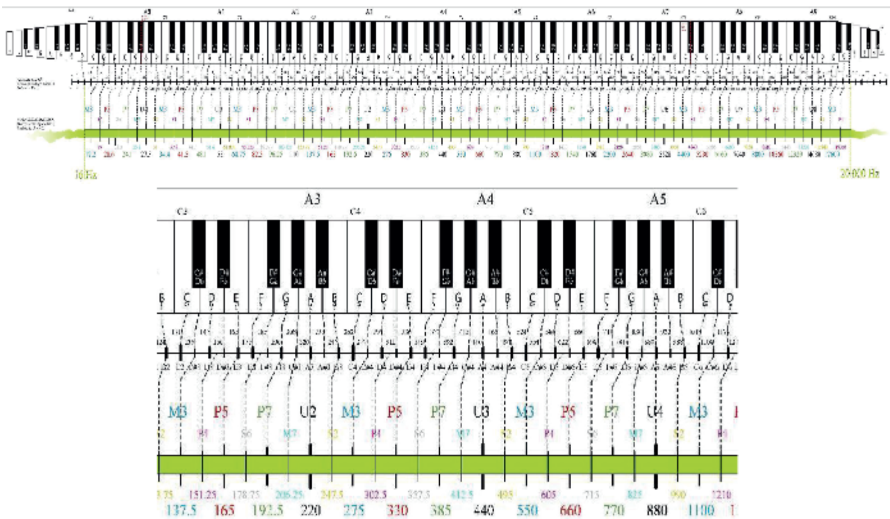

Figure 1.8. *Usual range of frequencies on a piano keyboard*