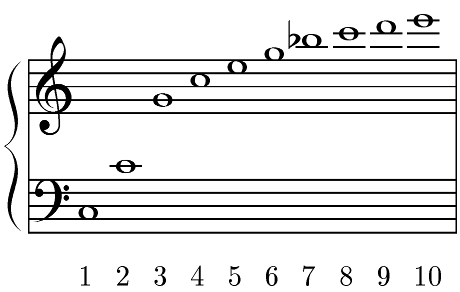
CHAPTER 4

Consonance and dissonance

In this chapter, we investigate the relationship between consonance and dissonance, and simple integer ratios of frequencies.

4.1. Harmonics

When a note on a stringed instrument or a wind instrument sounds at a certain pitch, say with frequency ν, all that really means is that the sound is (roughly) periodic with that frequency. The theory of Fourier series shows that such a sound can be decomposed as a sum of sine waves with various phases, at integer multiples of the frequency ν. The component of the sound with frequency ν is called the *fundamental*. The component with frequency $m\nu$ is called the mth *harmonic*, or the $(m-1)$st *overtone*. So for example if $m = 3$ we obtain the third harmonic, or the second overtone.[1]



This diagram represents the series of harmonics based on a fundamental at the C below middle C. The seventh harmonic is actually somewhat flatter than the B♭ above the treble clef. In the modern equally tempered scale, even the third and fifth harmonics are very slightly different from the notes G and E shown above—this is more extensively discussed in Chapter 5.

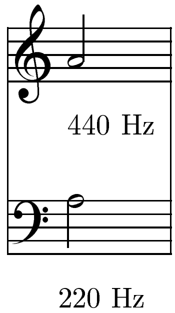
There is another word which we have been using in this context: the mth *partial* of a sound is the mth frequency component, counted from the bottom. So for example on a clarinet, where only the odd harmonics are present, the first partial is the fundamental, or first harmonic, and the second partial is the third harmonic. This term is very useful when discussing sounds where the partials are not simple multiples of the fundamental, such as for example the drum, the gong, or the various instruments of the gamelan.

Exercises

1. Define the following terms, making the distinctions between them clear:
(a) the mth harmonic, (b) the mth overtone, (c) the mth partial.

[1] I find that the numbering of overtones is confusing, and I shall not use this numbering.

4.2. Simple integer ratios



Why is it that two notes an octave apart sound consonant, while two notes a little more or a little less than an octave apart sound dissonant? An interval of one octave corresponds to doubling the frequency of the vibration. So for example, the A above middle C corresponds to a frequency of 440 Hz, while the A below middle C corresponds to a frequency of 220 Hz.

We have seen in Chapter 3 that if we play these notes on conventional stringed or wind (but not percussive) instruments, each note will contain not only a component at the given frequency, but also partials corresponding to multiples of that frequency. So for these two notes we have partials at:

440 Hz, 880 Hz, 1320 Hz, 1760 Hz, ...

220 Hz, 440 Hz, 660 Hz, 880 Hz, ...

On the other hand, if we play two notes with frequencies 440Hz and 225Hz, then the partials occur at:

440 Hz, 880 Hz, 1320 Hz, 1760 Hz, ...

225 Hz, 450 Hz, 675 Hz, 900 Hz, ...

The presence of components at 440 Hz and 450 Hz causes a sensation of roughness, which is interpreted by the ear as dissonance. We shall discuss at length, later in this chapter, the history of different explanations of consonance and dissonance, and why this should be taken to be the correct one.

Because of the extreme consonance of an interval of an octave, and its role in the series of partials of a note, the human brain often perceives two notes an octave apart as being "really" the same note but higher. This is so heavily reinforced by musical usage in every genre that we have difficulty imagining that it could be otherwise. When choirs sing "in unison", this usually means that the men and women are singing an octave apart.[2] The idea that notes differing by a whole number of octaves should be considered as equivalent is often referred to as *octave equivalence.*

The musical interval of a perfect fifth[3] corresponds to a frequency ratio of 3:2. If two notes are played with a frequency ratio of 3:2, then the third partial of the lower note will coincide with the second partial of the upper note, and the notes will have a number of higher partials in common. If, on the other hand, the ratio is slightly different from 3:2, then there will be a

[2]It is interesting to speculate what effect it would have on the theory of color if visible light had a span greater than an octave; in other words, if there were to exist two visible colors, one of which had exactly twice the frequency of the other. In fact, the span of human vision is just shy of an octave. This may explain why the colors of the rainbow seem to join up into a circle.

[3]We shall see in the next chapter that the fifth from C to G in the modern Western scale is not precisely a perfect fifth.

sensation of roughness between the third partial of the lower note and the second partial of the upper note, and the notes will sound dissonant.

The Experiences of Pythagoras
(Gaffurius, 1492)

In this manner, small integer ratios of frequencies are picked out as more consonant than other intervals. We stress that this discussion only works for notes whose partials are at multiples of the fundamental frequency. Pythagoras essentially discovered this in the sixth century B.C.; he discovered that when two similar strings under the same tension are sounded together, they give a pleasant sound if the lengths of the strings are in the ratio of two small integers. This was the first known example of a law of nature ruled by the arithmetic of integers, and greatly influenced the intellectual development of his followers, the Pythagoreans. They considered that a liberal education consisted of the "quadrivium", or four divisions: numbers in the abstract, numbers applied to music, geometry, and astronomy. They expected that the motions of the planets would be governed by the arithmetic of ratios of small integers in a similar way. This belief has become encoded in the phrase "the music of the spheres",[4] literally denoting the inaudible sound produced by the motion of the planets, and has almost disappeared in modern astronomy (but see the remarks in Exercise 1 of Section 6.2).[5]

4.3. Historical explanations of consonance

In writing this section, I have drawn heavily on the work of Plomp and Levelt. The reference can be found at the end of the section.

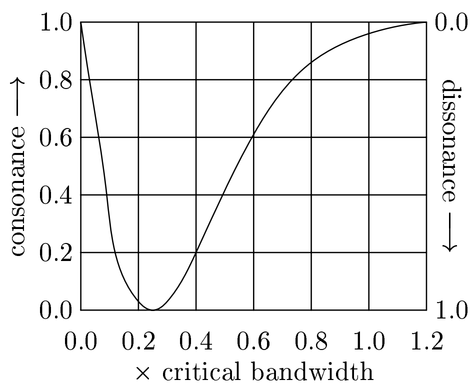
The discovery of the relationship between musical pitch and frequency occurred around the sixteenth or seventeenth century, with the work of

[4]Plato, *Republic*, 10.617, ca. 380 B.C.

[5]The idea embodied in the phrase "the music of the spheres" is still present in the seventeenth century work of Kepler on the motion of the planets. He called his third law the "harmonic law", and it is described in a work entitled *Harmonices Mundi* (Augsburg, 1619). However, his law properly belongs to physics, and states that the square of the period of a planetary orbit is proportional to the cube of the maximum diameter. It is hard to find any recognizable connection with musical harmony or the arithmetic of ratios of small integers. Kepler's ideas are celebrated in Paul Hindemith's opera, *Die Harmonie der Welt*, 1956–7. The title is a translation of Kepler's.

Galileo Galilei and (independently) Mersenne. Galileo's explanation of consonance was that if two notes have their frequencies in a simple integer ratio, then there is a regularity, or periodicity to the total waveform, not present with other frequency ratios, so that the ear drum is not "kept in perpetual torment".[6] The problem with this explanation is that it involves some circular reasoning—the notes are consonant because the ear finds them consonant! Furthermore, experimentation with tones produced using nonharmonic partials produce results which contradict this explanation, as we shall see in §4.6.

In the seventeenth century, it was discovered that a simple note from a conventional stringed or wind instrument had partials at integer multiples of the fundamental. The eighteenth century theoretician and musician Rameau ([**99**], chapter 3) regarded this as already being enough explanation for the consonance of these intervals, but Sorge[7] (1703–1778) was the first to consider roughness caused by close partials as the explanation of dissonance. It was not until the nineteenth century that Helmholtz (1821–1894) [**48**] sought to explain consonance and dissonance on a more scientific basis. Helmholtz based his studies on the structure of the human ear. His idea was that for small differences between the frequencies of partials, beats can be heard, whereas for larger frequency differences, this turns into roughness. He claimed that for maximum roughness, the difference between the two frequencies should be 30–40 Hz, independently of the individual frequencies. For larger frequency differences, the sense of roughness disappears and consonance resumes. He then goes on to deduce that the octave is consonant because all the partials of the higher note are among the partials of the lower note, and no roughness occurs.



Plomp and Levelt, in the nineteen sixties, seem to have been the first to carry out a thorough experimental analysis of consonance and dissonance for a variety of subjects, with pure sine waves, and at a variety of pitches. The results of their experiments showed that on a subjective scale of consonance ranging from zero (dissonant) to one (consonant), the variation with frequency ratio has the shape shown in the graph to the left. The x axis of this graph is labeled in multiples of the critical bandwidth, defined below. This means that the actual scale in Hertz on the horizontal axis of the graph varies according to the pitch of the notes, but the

[6] Galileo Galilei, *Discorsi e dimonstrazioni mathematiche interno à due nuove scienze attenenti alla mecanica ed i movimenti locali*, Elsevier, 1638. Translated by H. Crew and A. de Salvio as *Dialogues concerning two new sciences*, McGraw-Hill, 1963.

[7] G. A. Sorge, *Vorgemach der musicalischen Composition*, Verlag des Autoris, Lobenstein, 1745–1747

shape of the graph remains constant; the scaling factor was shown by Plomp and Levelt to be proportional to critical bandwidth.

The salient features of the above graph are that the maximum dissonance occurs at roughly one quarter of a critical bandwidth, and consonance levels off at roughly one critical bandwidth.

It should be stressed that this curve is for pure sine waves, with no harmonics; also that consonance and dissonance is different from recognition of intervals. Anyone with any musical training can recognize an interval of an octave or a fifth, but for pure sine waves, these intervals sound no more nor less consonant than nearby frequency ratios.

Exercises

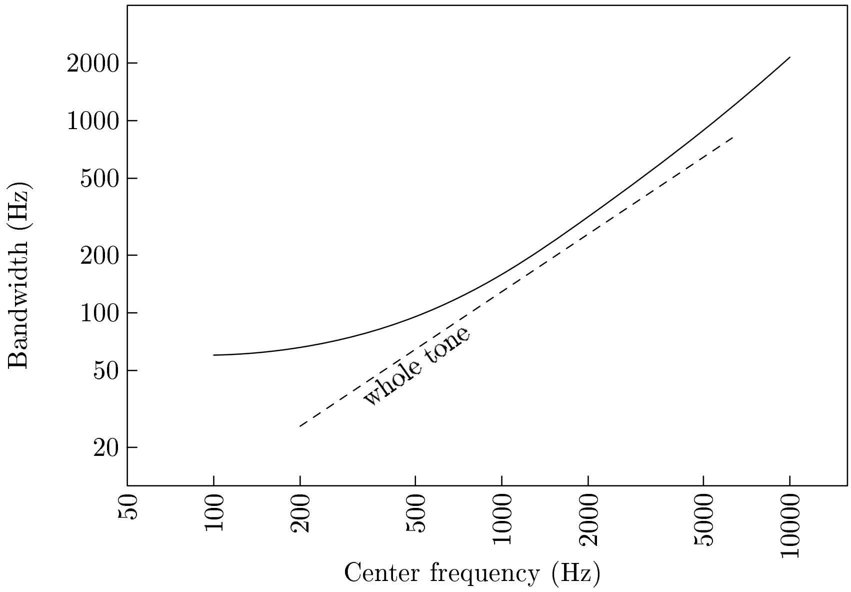
1. Show that the function $f(t) = A\sin(at) + B\sin(bt)$ is periodic when the ratio of a to b is a rational number, and nonperiodic if the ratio is irrational. [Hint: Differentiate twice and take linear combinations to get a single sine wave, to get information about possible periods]

Further reading:

R. Plomp and W. J. M. Levelt, *Tonal consonance and critical bandwidth*, J. Acoust. Soc. Am. 38 (1965), 548–560.

4.4. Critical bandwidth

To introduce the notion of *critical bandwidth*, each point of the basilar membrane in the cochlea is thought of as a band pass filter, which lets through frequencies in a certain band, and blocks out frequencies outside that band. The actual shape of the filter is certainly more complicated than this simplified model, in which the left, top and right edges of the envelope of the filter are straight vertical and horizontal lines. This is exactly analogous to the definition of bandwidth given in §1.10, and introducing a smoother shape for the filter does not significantly alter the discussion. The width of the filter in this model is called the critical bandwidth. Experimental data for the critical bandwidth as a function of center frequency is available from a number of sources, listed at the end of this section. Here is a rough sketch of the results.



Critical bandwidth as a function of center frequency

A rough calculation based on this graph shows that the size of the critical bandwidth is somewhere between a whole tone and a minor third throughout most of the audible range, and increasing to a major third for small frequencies.

Further reading:

B. R. Glasberg and B. C. J. Moore, *Derivation of auditory filter shapes from notched-noise data*, Hear. Res. 47 (1990), 103–138.

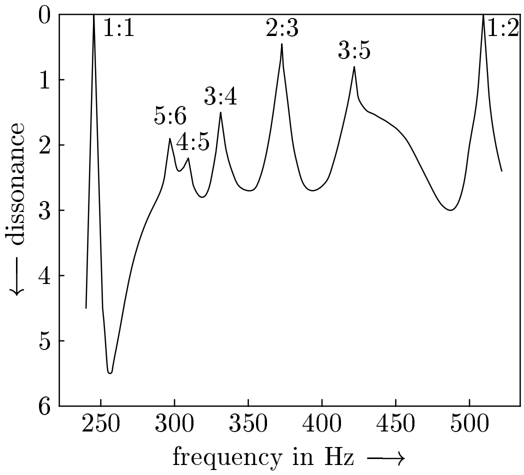
E. Zwicker, *Subdivision of the audible frequency range into critical bands (Frequenzgruppen)*, J. Acoust. Soc. Am. 33 (1961), 248.

E. Zwicker, G. Flottorp and S. S. Stevens, *Critical band width in loudness summation*, J. Acoust. Soc. Am. 29 (1957), 548–557.

E. Zwicker and E. Terhardt, *Analytical expressions for critical-band rate and critical bandwidth as a function of frequency*, J. Acoust. Soc. Am. 68 (1980), 1523–1525.

4.5. Complex tones

Plomp and Levelt took the analysis one stage further, and examined what would happen for tones with a more complicated harmonic content. They worked under the simplifying assumption that the total dissonance is the sum of the dissonances caused by each pair of adjacent partials, and



used the above graph for the individual dissonances. They do a sample calculation in which a note has partials at the fundamental and its multiples up to the sixth harmonic. The graph they obtain is shown to the right. Notice the sharp peaks at the fundamental (1:1), the octave (1:2) and the perfect fifth (2:3), and the smaller peaks at ratios 5:6 (just minor third), 4:5 (just major third), 3:4 (perfect fourth) and 3:5 (just major sixth). If higher harmonics are taken into account, the graph acquires more peaks.

In order to be able to draw such Plomp–Levelt curves more systematically, we choose a formula which gives a reasonable approximation to the curve displayed on page 106. Writing x for the frequency difference in multiples of the critical bandwidth, we choose the dissonance function to be[8]

$$f(x) = 4|x|e^{1-4|x|}.$$

This takes its maximum value $f(x) = 1$ when $x = \frac{1}{4}$, as can easily be seen by differentiating. It satisfies $f(0) = 0$, and $f(1)$ is small (about $\frac{1}{5}$), but not zero. This last feature does not quite match the graph given by Plomp and Levelt, but a closer examination of their data shows that the value $f(1) = 0$ is not quite justified.

Further reading:

R. Plomp and W. J. M. Levelt, *Tonal consonance and critical bandwidth*, J. Acoust. Soc. Am. 38 (1965), 548–560.

4.6. Artificial spectra

So what would happen if we artificially manufacture a note having partials which are not exact multiples of the fundamental? It is easy to perform such experiments using a digital synthesizer. We make a note whose partials are at

440 Hz, 860 Hz, 1203 Hz, 1683 Hz, ...

and another with partials at

225 Hz, 440 Hz, 615 Hz, 860 Hz, ...

to represent slightly squeezed harmonics. These notes sound consonant, despite the fact that they are slightly less than an octave apart, whereas scaling the second down to

220 Hz, 430 Hz, 602 Hz, 841 Hz, ...

[8]Sethares [**119**] takes for the dissonance function $f(x) = e^{-b_1 x} - e^{-b_2 x}$ where $b_1 = 3.5$ and $b_2 = 5.75$. This needs normalizing by multiplication by about 5.5, and then gives a graph very similar to the one I have chosen. The particular choice of function is somewhat arbitrary, because of a lack of precision in the data as well as in the subjective definition of dissonance. The main point is to mimic the visible features of the graph.

causes a distinctly dissonant sounding exact octave.

If we are allowed to change the harmonic content of a note in this way, we can make almost any set of intervals seem consonant. This idea was put forward by Pierce (1966, reference below), who designed a spectrum suitable for an equal temperament scale with eight notes to the octave. Namely, he used the following partials, given as multiples of the fundamental frequency:

$$1:1, \quad 2^{\frac{5}{4}}:1, \quad 4:1, \quad 2^{\frac{5}{2}}:1, \quad 2^{\frac{11}{4}}:1, \quad 8:1.$$

This may be thought of as a stretched version of the ordinary series of harmonics of the fundamental. When two notes of the eight tone equal tempered scale are played using synthesized tones with the above set of partials, what happens is that the partials either coincide or are separated by at least $\frac{1}{8}$ of an octave. Pierce's conclusion is that

> ...by providing music with tones that have accurately specified but nonharmonic partial structures, the digital computer can release music from the tyrrany of 12 tones without throwing consonance overboard.

Further reading:

W. Hutchinson and L. Knopoff, *The acoustic component of western consonance*, Interface 7 (1978), 1–29.

A. Kameoka and M. Kuriyagawa, *Consonance theory I: consonance of dyads*, J. Acoust. Soc. Am. 45 (6) (1969), 1452–1459.

A. Kameoka and M. Kuriyagawa, *Consonance theory II: consonance of complex tones and its calculation method*, J. Acoust. Soc. Am. 45 (6) (1969), 1460–1469.

Jenö Keuler, *Problems of shape and background in sounds with inharmonic spectra*, Music, Gestalt, and Computing [**65**], 214–224, with examples from the accompanying CD.

Max V. Mathews and John R. Pierce, *Harmony and nonharmonic partials*, J. Acoust. Soc. Am. 68 (1980), 1252–1257.

John R. Pierce, *Attaining consonance in arbitrary scales*, J. Acoust. Soc. Am. 40 (1966), 249.

John R. Pierce, *Periodicity and pitch perception*, J. Acoust. Soc. Am. 90 (4) (1991), 1889–1893.

W. A. Sethares, *Tuning, timbre, spectrum, scale* [**119**]. This book comes with a compact disc full of illustrative examples.

W. A. Sethares, *Consonance-based spectral mappings.* Computer Music Journal 22 (1) (1998), 56–72.

Frank H. Slaymaker, *Chords from tones having stretched partials.* J. Acoust. Soc. Am. 47 (1970), 1569–1571.

E. Terhardt, *Pitch, consonance, and harmony.* J. Acoust. Soc. Am. 55 (1974), 1061–1069.

E. Terhardt and M. Zick, *Evaluation of the tempered tone scale in normal, stretched, and contracted intonation.* Acustica 32 (1975), 268–274.

4.7. Combination tones

When two loud notes of different frequencies f_1 and f_2 are played together, a note can be heard corresponding to the difference $f_1 - f_2$ between the two frequencies. This was discovered by the German organist Sorge (1744) and Romieu (1753). Later (1754) the Italian violinist Tartini claimed to have made the same discovery as early as 1714. Helmholtz (1856) discovered that there is a second, weaker note corresponding to the sum of the two frequencies $f_1 + f_2$, but that it is much harder to perceive. The general name for these sum and difference tones is *combination tones*, and the difference notes in particular are sometimes called *Tartini's tones*. The reason (overlooked by Helmholtz) why the sum tone is so hard to perceive is because of the phenomenon of masking discussed at the end of §1.2.

It is tempting to suppose that the combination tones are a result of a discussion similar to the discussion of beats in §1.7. However, this seems to be misleading, as this argument would seem more likely to give rise to notes of half the difference and half the sum of the notes, and this does not seem to be what occurs in practice. Moreover, when we hear beats, we are not hearing a *sound* at the beat frequency, because there is no corresponding place on the basilar membrane for the excitation to occur. Further evidence that these are different phenomena is that when the two tones are heard one with each ear, beats are still discernable, while combination tones are not.

Helmholtz [**48**] (Appendix XII) had a more convincing explanation of combination tones, based on the supposition that the sounds are loud enough for nonlinearities in the response of some part of the auditory system to come into effect.

In the presence of a quadratic nonlinearity, a damped harmonic oscillator with a sum of two sinusoidal forcing terms of different frequencies will vibrate with not only the two incoming frequencies but also with components at twice these frequencies and at the sum and difference of the frequencies. Intuitively, this is because

$$(\sin mt + \sin nt)^2 = \sin^2 mt + 2 \sin mt \sin nt + \sin^2 nt$$
$$= \tfrac{1}{2}(1 - \cos 2mt) + \tfrac{1}{2}(\cos(m-n)t - \cos(m+n)t) + \tfrac{1}{2}(1 - \cos 2nt).$$

So if some part of the auditory system is behaving in a nonlinear fashion, a quadratic nonlinearity would correspond to the perception of doubles of the incoming frequencies, which are probably not noticed because they look like overtones, as well as sum and difference tones corresponding to the terms $\cos(m+n)t$ and $\cos(m-n)t$.

Quadratic nonlinearities involve an asymmetry in the vibrating system, whereas cubic nonlinearities do not have this property. So it seems reasonable to suppose that the cubic nonlinearities are more pronounced in effect than the quadratic ones in parts of the auditory system. This would mean that combination tones corresponding to $2f_1 - f_2$ and $2f_2 - f_1$ would be more prominent than the sum and difference. This seems to correspond to what is experienced in practice. These cubic terms can be heard even at low volume, while a relatively high volume is necessary in order to experience the sum and difference tones.

Helmholtz's theory ([**48**], appendix XII) was that the nonlinearity giving rise to the distortion was occurring in the middle ear, and in particular the tympanic membrane. Measurements made by Guinan and Peake[9] have shown that the nonlinearities in the middle ear are insufficient to explain the phenomenon. Current theory favors an intracochlear origin for the nonlinearities responsible for the sum and difference tone. Furthermore, the distortions responsible for cubic effects are now thought to have their origins in psychophysical feedback, and are part of the normal auditory function rather than a result of overload.[10]

There is also a related concept of *virtual pitch* for a complex tone. If a tone has a complicated set of partials, we seem to assign a pitch to a composite tone by very complicated methods which are not well understood. Schouten[11] demonstrated that Helmholtz's discussion does not completely explain what happens for these more complex sounds. If the ear is simultaneously subjected to sounds of frequencies 1800 Hz, 2000 Hz and 2200 Hz then the subject hears a tone at 200 Hz, representing a "missing fundamental," and which might be interpreted as a combination tone. However, if the sounds have frequencies 1840 Hz, 2040 Hz and 2240 Hz then instead of hearing a 200 Hz tone as would be expected by Helmholtz's theory, the subject actually hears a tone at 204 Hz. Schouten's explanation for this has been disputed in more recent work, and it is probably fair to say that the subject is still not well understood.

Walliser[12] has given a recipe for determining the perceived missing fundamental, without supplying a mechanism which explains it. His recipe consists of determining the difference in frequency between two adjacent partials (or harmonic components of the sound), and then approximating this with as simple as possible a rational multiple of the lowest harmonic component. So in the above example, the difference is 200 Hz, so we take one nineth of 1840 Hz to give a missing fundamental of 204.4 Hz. This is an extremely good approximation to what is actually heard. Later authors have proposed minor modifications to Walliser's algorithm, for example by replacing the lowest partial with the most "dominant" in a suitable sense. A more detailed discussion can be found in chapter 5 of B. C. J. Moore's book [**79**].

---

[9] J. J. Guinan and W. T. Peake, *Middle ear characteristics of anesthetized cats.* J. Acoust. Soc. Am. **41** (1967), 1237–1261.

[10] See for example Pickles [**93**], pp. 107–109.

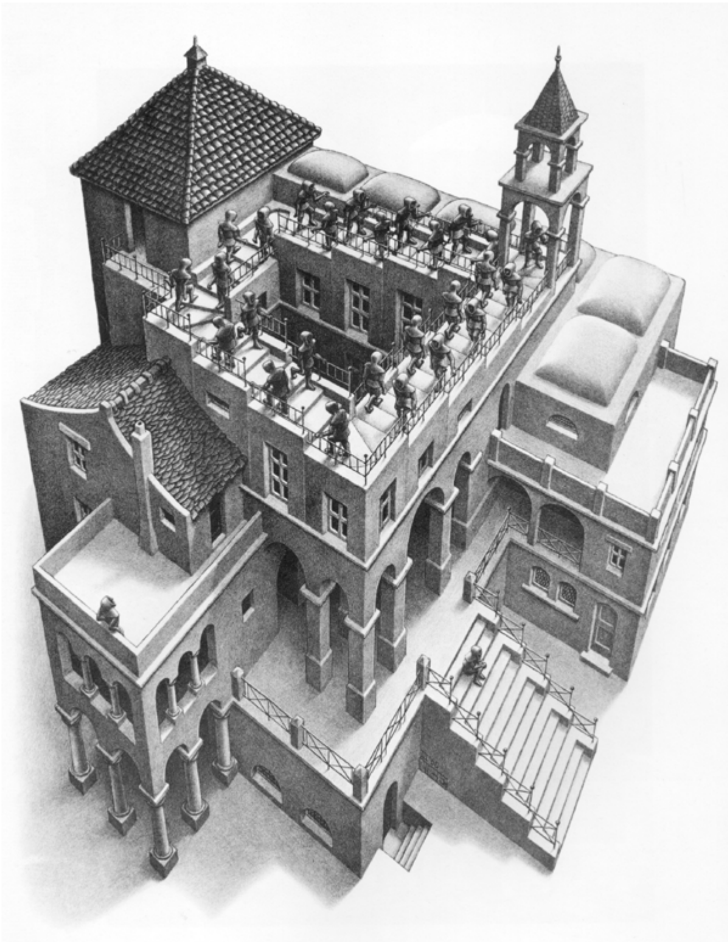
[11] J. F. Schouten, *The residue and the mechanism of hearing*, Proceedings of the Koningklijke Nederlandse Akademie van Wetenschappen **43** (1940), 991–999.

[12] K. Walliser, *Über ein Funktionsschema für die Bildung der Periodentonhöhe aus dem Schallreiz*, Kybernetik **6** (1969), 65–72.

Licklider[13] also cast doubt on Helmholtz's explanation for combination tones by showing that a difference tone cannot in practice be masked by a noise with nearby frequency, while it should be masked if Helmholtz's theory were correct.

Combination tones and virtual pitch remain among many interesting topics of modern psychoacoustics, and a current active area of research.

4.8. Musical paradoxes

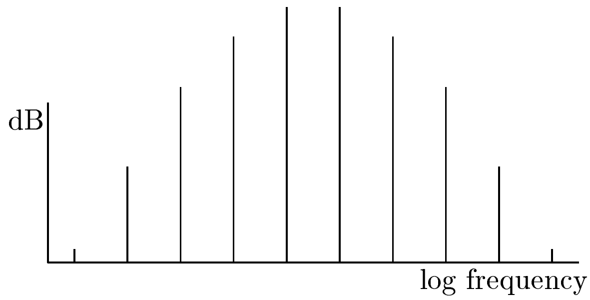


M. C. Escher, *Ascending and descending* (1960).

One of the most famous paradoxes of musical perception was discovered by R. N. Shepard, and goes under the name of the Shepard scale. Listening to the Shepard scale, one has the impression of an ever-ascending scale where the end joins up with the beginning, just like Escher's famous ever ascending staircase in his picture, *Ascending and descending*. This effect is achieved by building up each note out of a complex tone consisting of

[13] J. C. R. Licklider, *Periodicity by "pitch" and place "pitch"*, J. Acoust. Soc. Am. **26**, (1954), 945.

ten partials spaced at one octave intervals. These are passed through a filter so that the middle partials are the loudest, and they tail off at both the bottom and the top. The same filter is applied for all notes of the scale, so that after ascending through one octave, the dominant part of the sound has shifted downwards by one partial.



The partials present in this sound are of the form $2^n.f$, where f is the lowest audible frequency component.

A related paradox, discovered by Diana Deutsch, is called the *tritone paradox*. If two Shepard tones are separated by exactly half an octave (a tritone in the equal tempered scale), or a factor of $\sqrt{2}$, then it might be expected that the listener would be confused as to whether the interval is ascending or descending. In fact, only some listeners experience confusion. Others are quite definite as to whether the interval is ascending or descending, and consistently judge half the possible cases as ascending and the complementary half as descending.

Further reading:

E. M. Burns, *Circularity in relative pitch judgments: the Shepard demonstration revisited, again*, Perception and Psychophys. 21 (1977), 563–568.

D. Deutsch, *Musical illusions*, Scientific American 233 (1975), 92–104.

D. Deutsch, *A musical paradox*, Music Percept. 3 (1986), 275–280.

D. Deutsch, *The tritone paradox: An influence of language on music perception*, Music Percept. 8 (1990), 335–347.

R. N. Shepard, *Circularity in judgments of relative pitch*, J. Acoust. Soc. Am. 46 (1960), 2346–2353.

Further listening: (See Appendix R)

Auditory demonstrations CD (Houtsma, Rossing and Wagenaars), track 52 is a demonstration of Shepard's scale, followed by an analogous continuously varying tone devised by Jean-Claude Risset.