# Sound Design of Machines From a Musical Perspective

Hugo Fastl
AG Technische Akustik, MMK, TU München,
Arcisstr. 21, D-80333 München, Germany
E-mail: Fastl@mmk.ei.tum.de

## Introduction

In this article, relations between sound quality engineering and musical acoustics are presented. First, the concept of musical consonance proposed by Terhardt, based on the seminal classic work of Helmholtz is discussed. In particular, the view of Helmholtz on modern psychoacoustic magnitudes important in sound quality design such as fluctuation strength and roughness are illustrated.

The application of Terhardt's concept of sensory pleasantness in sound quality engineering is discussed, and advantages and disadvantages are shown. The concept of virtual pitch, i.e., a pitch sensation, which has no direct correspondence in the spectrum, is traced back to the concept of "basse fondamentale" put forward by Rameau in the 18th century. Characteristics of musical keys as advocated by Kirnberger and Schubart are displayed in view of the sound quality of the interior sound of passenger cars. Musical dynamics from *ppp* to *fff* are contrasted to current category scales for loudness scaling and the relations of loudness and tone color. Finally, examples of audio-visual interactions in speech and music as well as sound quality rating are given.

## The Concept of Musical Consonance

The concept of musical consonance as advocated by Terhardt[1] is illustrated in Fig. 1. Musical consonance depends on two features, namely harmony and sensory consonance. The concept of harmony was developed for musical sounds, whereas the concept of sensory consonance can be applied not only to musical sounds, but to all categories of sounds, e.g., sounds usual in sound quality engineering. Nevertheless, the concept of harmony also can have an impact with respect to sound quality. A basic result of the work of Helmholtz[2] was his discovery that the hearing sensation called roughness strongly influences sensory consonance. In modern terminology (e.g., Zwicker and Fastl,[3]), the concept of Helmholtz would be expressed as fluctuation strength and roughness being mainly based on interactions of spectral components of sounds. Figure 2 gives an illustration of the concept of Helmholtz in musical notation. In Fig. 2, the unfilled notes represent the sounds played, e.g., on a piano. The filled notes indicate harmonics. In example 1, the musical interval of an octave is displayed. The sounds played are $c_2$ and $c_3$ as indicated by unfilled notes. In addition to $c_2$ the harmonic $c_3$ (filled note) is audible, which gives a perfect match to the presented (unfilled) note, $c_3$. This means that in the musical concept, the octave $c_2$-$c_3$ represents a perfect consonance without beats because the second harmonic of $c_2$ ($c_3$) is identical to the higher note $c_3$. In contrast, example 2, which consists of the notes $c_2$ and $b_2$, represents a dissonance because the presented note, $b_2$, is a semitone apart from the second harmonic, $c_3$, of the note $c_2$. A similar argument holds for example 3.

Examples 4 and 5 of Helmholtz again show consonant intervals, which is easily seen in example 5. The notes played are $d_4$ and $a_4$ which form the musical interval of a fifth. The second harmonic of $d_4$ is $d_5$, and the third harmonic $a_5$, whereas the second harmonic of $a_4$ is $a_5$. Since the second harmonic of $a_4$ and the third harmonic of $d_4$ coincide, the interval of a fifth is considered to be a consonance. This reasoning is described in more detail in figure 3 from Helmholtz, where he illustrates the roughness of different musical intervals. The intervals considered always start with $c_4$, and the magnitude of the peaks indicates the magnitude of the roughness of the respective interval. As is easily seen in Fig.3, the octave, namely the interval $c_4$-$c_5$, represents a perfect consonance and no roughness is visible. Also for the fifth, $c_4$-$g_4$, Helmholtz displays no roughness values. On the contrary, the roughness of musical intervals reaches, according to Helmholtz, a maximum for one musical semitone ($c_4$-$cis_4$) or for eleven semitones ($c_4$-$b_4$). The results, displayed in Fig.3, clearly show that the concept of musical consonance is based on the absence of the psychoacoustic magnitude roughness, which we use today in sound quality engineering. In line with classic music theory of

From: Noise/News International, 2004, Vol. 12/4

Pythagoras dating back to 500 BC, the musical intervals of octave, fifth and fourth with frequency ratios of 1:2, 2:3, and 3:4 are considered as perfect consonances because they show no fluctuation strength or roughness.

Based on the classic findings by Helmholtz[2], Terhardt and Stoll[4] proposed the concept of *sensory pleasantness*, which is illustrated in Fig. 4. As expected from the classic works of music theory, sensory pleasantness decreases with increasing roughness (Fig 4a). Also, in line with data from Helmholtz[2] , Fig 4b shows that sensory pleasantness decreases with increasing sharpness. Figures 4c and 4d indicate that sensory pleasantness increases with increasing tonality, but decreases with increasing loudness.

While the results displayed in Figs. 4a, 4b, and 4d are well known in sound quality engineering and sound quality design, the results displayed in Fig. 4c must be considered in more detail. It is clear that in a musical context, as displayed in Fig. 1, sensory consonance and hence sensory pleasantness increases with the tonal character of a sound. However, in noise control engineering, exactly the opposite behavior may occur! In many standards, tonal components of sounds are "punished" by tone penalties. For example, in the German standard for noise immissions from industrial noise sources (TA Lärm), for sounds with clearly audible tonal components, a tone correction of up to 6 dB is added to the measured A-weighted sound level. Calculations for tone corrections are also proposed in German standard DIN 45681. This means that, in a musical context, tonal components usually have a positive impact, whereas in noise control engineering generally tonal components should be avoided.

An interesting example in this context comes from Japan: A manufacturer of wire matrix printers controlled the sequence of the wires in such a way that the printer played well known tunes. When this feature was introduced, it was very welcome by the users because of its novelty. However, after few days it can be quite annoying if a printer plays the same tunes over and over again.

## Roots of Musical Chords and Virtual Pitch

It is well known in music theory, that musical chords represent a specific tonality, frequently called the "root" of the chord. This means that the "root" represents a suitable bass note and further can be responsible for the tonality of a piece of music, e.g., that a tune is written in C-major. As displayed in Fig. 1, this relationship belongs to the concept of harmony. It dates back to the 18th century and is described in great detail by the famous French music theorist, Rameau.[5] In musical terms, the notes displayed in Fig. 5 illustrate that the chord g c e has the "meaning" of C-major. In music theory, it is explained that the notes $g_4$, $c_5$, and $e_5$ are considered to be the third, fourth, and fifth harmonic of a common root. As shown in Fig. 5, the first harmonic, which fits the chord displayed by filled notes is the unfilled note $c_3$, and the second harmonic is the unfilled note $c_4$. The frequencies of these notes form the ratio 1:2:3:4:5 and hence correspond to the musical intervals octave, fifth, fourth, and major third. As we have seen, up to the fourth, these intervals have, since ancient times, been considered to be consonant. The major third received more and more acceptance as a consonant interval in the 17th century, but was a little suspect to older music theorists like Pythagoras.

Rameau's concept[5] of "basse fondamentale" has its modern counterpart in the virtual pitch theory of Terhardt,[6] which is illustrated in Fig. 6. In the example displayed in Fig. 6, a harmonic complex tone with a basic frequency of 200 Hz is considered, from which the first two harmonics (200 Hz and 400 Hz) are removed. In essence, the concept of virtual pitch predicts that from the spectral pitches of each harmonic, subharmonics are calculated by integer ratios (1:1, 1:2, 1:3, 1:4, etc.). The number of coincident calculated subharmonics gives an indication of the virtual pitch perceived. As becomes clear from the example given in Fig. 6, virtual pitch has some ambiguity. Terhardt's model calculates the virtual pitch of the incomplete complex harmonic tone near 200 Hz, but also pitches one octave lower or one octave higher are candidates. This ambiguity is in line with the musical concept shown in Fig. 5, because the "root" of the chord displayed is represented by $c_3$, but $c_4$ is also possible.

For sound quality engineering, the concept of virtual pitch, which is based on the musical concept of harmony, plays an important role: the sound of a product can produce a pitch sensation in a frequency region where no spectral components are present! When assessing sound quality by physical means like spectral analysis, it always has to be kept in mind that audible tonal components can be virtual pitches, which show up in the spectrum only by higher harmonics.

### Character of Musical Keys

It is a long-standing debate, whether musical keys can represent a specific character. In particular, Kirnberger[7] and Schubart[8] related different keys to different emotions or moods. Experiments by Kunkel[9] as well as our own experience showed that with equal temperament it is essentially impossible to correlate different emotions with different musical keys. However, in well tempered intonations, i.e., a tuning put forward in particular by Werckmeister[10] and used by the famous composer Johann Sebastian Bach, characteristics of musical keys may be perceived by experts.

Since a famous German manufacturer of passenger cars insists that a luxury car has to be tuned in minor and not in major, Fig. 7 gives (more as curiosity) the characteristics of minor keys in English—translated from the original German. The original German text is shown in Appendix A.

d a e g h c fis f cis b gis dis

| | | |
|---|---|---|
| First Category | A - minor | *Pious womanhood and softness of character* |
| (Womanhood and Gentleness) | E - minor | *Naïveté, the innocent declaration of love by a young girl, clothed in white, with a rose-red bow on the breast* |
| | B - minor | *Melancholic womanhood obsessing about something.* |
| | D - minor | *Sound of patience, of quiet waiting for their destiny, and of surrender to the divinely foreordained.* |
| Second Category | F# - minor | *Dark tone.* |
| (Oppressive sadness) | C# - minor | *Repentant lamentation, the sigh of unsatisfied friendship and love.* |
| | G# - minor | *A grumbler/misery, a heart pressed almost to suffocation, a lamentation that sighs into the Double Cross (?), a difficult battle. In short, the color of this key is everything that is very hard and laborious.* |
| | E*b* - minor | *A feeling of disquiet for all the deepest compulsions of the soul; of inwardly brooding despair, of blackest gloom, of bleakest morale.* |
| Third Category | G - minor | *Dissatisfaction, uneasiness, resentment, listlessness.* |
| (Hopeless sadness) | C - minor | *Lamentation of unhappy love.* |
| | F - minor | *Deep gloom, the lament of the dead, miserable moaning and the yearning for death.* |
| | B*b* - minor | *Moaning, reproaches against God and the world, preparation for suicide.* |

Figure 7. Characteristics of musical keys in minor according to Schubart (1806). English version of the original German given in Appendix A1. The (?) indicates difficulty in translating the exact sentiment. (Many thanks to Aaron Hastings, Rajani Ippili and Eckhard Groll, Purdue University, for this translation.)

From all the keys given in Fig. 7, only the first two namely a-minor and e-minor may be useful for sound design because all the other keys represent feelings of sadness and despair. However, according to the results displayed in Fig 7, a-minor should have a religious, feminine character, and e-minor should be related to the love of a naïve girl. Although this author has studied music, it is not easy to detect the character of these keys in a minor key. How to tune the interior sound of a car in a minor key, whether this is a-minor or e-minor, seems to be the secret of the German car manufacturer.

## Musical Dynamics

As is well known from music theory, musical dynamics range between very soft (pianissimo *pp*) and very loud (fortissimo *ff*). For extreme values of loudness the musical notations piano pianissimo (*ppp*) and forte fortissimo (*fff*) are used. Figure 8 enables a comparison of musical dynamics with category scales frequently used for sound quality engineering.

| ***pp*** | ***p*** | ***mf*** | ***f*** | ***ff*** |
|---|---|---|---|---|
| very soft | soft | neither soft or loud | loud | very loud |

| ***ppp*** | ***pp*** | ***p*** | ***mf*** | ***f*** | ***ff*** | ***fff*** |
|---|---|---|---|---|---|---|
| extremely soft | very soft | soft | neither soft or loud | loud | very loud | extremely loud |

Figure 8. Musical Dynamics and Category Scales.

The upper panel shows a five-step category scale, the lower panel a seven-step category scale. It is interesting to note that there is good correspondence between the symbols of musical dynamics and loudness rating. However, for the middle category (mezzoforte, *mf,* in musical notation), the interpretation is a little different between sound quality engineering and musical acoustics. In sound engineering, the category neither soft nor loud is just in the middle between the category soft on the one hand, and the category loud on the other hand. In musical notation, however, meozzforte means a little loud or half loud, and therefore is not centered between soft and loud, but closer to loud. Despite these small differences, category scalings with five step scales or seven step scales have proven successful in sound quality engineering.

When discussing loudness, two important influences for musical acoustics as well as sound engineering must be discussed. As an example, Fig. 9 shows the loudness patterns for the musical note $f_4$ played by a horn in a reverberant chamber at the musical dynamics pianissimo, *pp,* and fortissimo, *ff.* As becomes clear from the loudness values N in sones as well as the areas of the loudness distributions, for fortissimo the loudness is about a factor of eight larger than the loudness for pianissimo. However, in addition, a considerable shift in spectral distribution can be seen. While for pianissimo the first harmonic near 4 Bark is dominant, for fortissimo the second harmonic near 8 Bark takes over. This is related to the blaring sound of a horn when played at fortissimo. Similar interactions of loudness and tone color are well known from sound quality engineering. Therefore, for modifications, not only the differences in level, but also differences in spectral distribution have to be taken into account. Loudness patterns as displayed in Fig. 9 represent a valuable tool to evaluate differences in loudness, and at the same time in tone color. Figure 10 again illustrates that at same level, for sounds with different spectral distribution, the perceived loudness can differ considerably.

For the same A-frequency weighted level of 85 dB, the left panel shows the loudness pattern of a piccolo flute, the right panel of a pipe organ with full registers. Because of the larger bandwidth of the organ sound, at same level, the perceived loudness is about a factor of two larger. These effects of bandwidth of the sound source are well known in music and assessed by conductors as well as sound recording engineers. They also can play an important part in sound quality engineering and design.

## Audio-Visual Interactions

The last example illustrates audiovisual interactions, which can play a role both in musical acoustics and sound quality engineering. Figure 11 shows a simplified floor plan of a concert hall and its stage. A speaker S is placed on the stage, and speech is recorded at different postions—1, 2, and 3—in the concert hall.

The rating of speech quality is done either by just presenting the sound recorded at positions 1, 2, and 3, or by additional presentation of the correlated visual image. The corresponding ratings of speech quality are given in Fig 12. Unfilled symbols represent ratings with acoustic presentation only; filled symbols indicate ratings for acoustical plus visual presentation.

The data displayed in Fig. 12 show that for the acoustical presentation alone, speech quality is rated to be fair. If, however, in addition to the acoustical presentation a photo of the situation in the concert hall is presented to the subjects, despite identical acoustical input, the speech quality rating may be influenced. At position 1, with the additional presentation of the visual image, the rating gets poorer. At position 2, the visual image has no influence on the rating, whereas at position 3, the rating improves with the visual image. These results may be interpreted as follows. In position 1, the source is very close to the receiver, and therefore with the visual image the subject expects a very clear sound, which does not occur for speech in a concert hall with a reverberation time of about two seconds. At position 2, the visual image shows a larger distance to the source, and therefore the subjects seem to accept some compromise with respect to sound quality. Interestingly, at position 3, the presentation of the visual image improves the rating. This may be due to the fact that in the visual image, the subject realizes a large distance between source and receiver, and feels that—given that adverse situation—the speech quality can be considered as good.

Another effect of audio-visual interaction is well known to sound recording engineers. If a musician inadvertently hits his music stand with his instrument, the audience in the concert hall hardly notices the related impulsive sound. However, when recorded on a CD, the same noise is definitely unacceptable. Because, in the concert hall, the visual information can explain the reason for the noise, it remains largely unnoticed.

Similar audio-visual interactions were reported by Suzuki *et al.*[11] with respect to the sound quality of noises. They showed that when the sound of white noise is combined with the image of a waterfall, the rating of the sound quality improves. Likewise, Hashimoto and Hatano[12] demonstrated that visual images can influence the rating of car sounds considerably. In an experiment with semantic differential, they measured the rating of the interior sound of passenger cars for acoustical presentation alone, or for the combination of an acoustical signal plus visual presentation of the moving car. Compared with the acoustical presentation alone, the rating of the adjectives “pleasant” or “quiet” improved by one step of the seven-step scale with the presentation of the corresponding visual image. For many adjective pairs, the improvement in rating with the addition of the visual image corresponded to a decrease in sound pressure level by as much as 10 dB! Patsouras and Böhm[13] reported that the influence of a visual input on the rating of interior sounds of cars increased with the complexity of the visual presentation. In addition, Patsouras *et al.*[14] reported that the loudness rating for train noise may also be influenced by the additional presentation of visual images.

((is it better to place this just at the end of the conlusion??))
References to a large number of papers on noise, psychoacoustics, and music may be found on the Internet.[16,17]

## Conclusion

The few examples described in this paper indicate that sound quality engineering and musical acoustics have several aspects in common. In addition, it is interesting to note that tools used today we use in sound quality engineering, were first developed for musical acoustics. For example, dummy head recordings were used in the 1960s for evaluations of the sound quality of concert halls. Despite the fact that dummy heads still are used in a musical context, the majority of today's instruments are used in the context of sound quality engineering. On the other hand, the analysis of radiation patterns has gone back and forth between musical acoustics and sound quality design. In the 18th century, Chladni[15] illustrated the radiation patterns of musical instruments by putting small seeds on them, showing the distribution of regions with large or small vibration. These days, laser vibrometers, which enable a very detailed analysis of the vibration pattern of products, are very helpful tools for sound quality engineering. However, they are also used to measure the vibration patterns of musical instruments. Despite all of this modern equipment, the mystery e.g., of the superb sound quality of old Italian violins has not yet been revealed. This means that although we have achieved a high standard both in sound quality design and musical acoustics, many open questions still wait to be solved.

## Acknowledgements

The author wishes to thank Dipl.-Ing. Christine Patsouras for elucidating discussions on well tempered musical scales and characteristics of musical keys. The editorial assistance by her and Dipl.-Ing. Melanie Böhm is gratefully acknowledged. Part of this work is supported by Deutsche Forschungsgemeinschaft, FA 140/2.

## References


1. Terhardt, E., The concept of musical consonance: A link between music and psychoacoustics. *Music Perception,* **1,** 276-295 (1984).
2. Helmholtz, A., Die Lehre von den Tonempfindungen. Vieweg Braunschweig (1863).
3. Zwicker, E., Fastl, H., Psychoacoustics-Facts and Models. 2nd Ed. Springer Heidelberg (1999)
4. Terhardt, E., Stoll, G., Skalierung des Wohlklangs (der sensorischen Konsonanz) von 17 Umweltschallen und Untersuchung der beteiligten Hörparameter. *Acustica,* **48,** 247-253 (1981).
5. Rameau, J. P., Démonstration du principe de l'harmonie servant de basse à tout l'art musical théoretique et pratique, Paris (1750).
6. Terhardt, E., Pitch, consonance, and harmony. *J. Acoust. Soc. Am.* **55,** 1061-1069 (1974).
7. Kirnberger, J. P., Gedanken über die verschiedenen Lehrarten in der Composition (1782).
8. Schubart, Ch., Ideen zu einer Ästhetik der Tonkunst (1806).
9. Kunkel, F. J., Das Tonsystem in Zahlen (1877).
10. Werckmeister, A., Harmonologia musica oder kurtze Anleitung zur musicalischen Composition (1702).
11. Suzuki, Y., Abe, K., Ozawa, K., Sone, T., Factors for perceiving sound environments and the effects of visual and verbal information on these factors. In: Contributions to Psychological Acoustics (A. Schick et al. Eds), BIS Oldenburg, 209-232 (2000).
12. Hashimoto, T., Hatano, S., Effects of Factors other than Sound to the Perception of Sound Quality. In: Proc. 17th ICA (2001).
13. Patsouras, Ch., Böhm, M. Visueller Kanal oder Ablenkung? Einfluß einer optischen Komponente auf die Beurteilung der Geräuschqualität von Fahrzeug-Innengeräuschen. In: Fortschritte der Akustik, DAGA 02, DEGA Oldenburg (2002)
14. Patsouras, Ch., Filippou, Th., D., Fastl, H., Influences of color on the loudness judgement. In: Proc. FORUM ACUSTICUM, Sevilla (2002).
15. Chladni, E., Entdeckungen über die Theorie des Klanges. Leipzig (1787).
16. http://www.mmk.ei.tum.de/music.html
17. http://www.mmk.ei.tum.de/noise.html

**Appendix A. Original German version of Figure 7.**

d a e g h c fis f cis b gis dis

## Characteristics of Minor Keys by Schubart

| | | |
|---|---|---|
| 1. Klasse | a - Moll | *Fromme Weiblichkeit unt Weichheit des Charakters* |
| (Weiblichkeit unt Sanftmut) | e - Moll | *naive, weibliche unschuldige Liebeserklärung vergleichbar mit einem Mädchen, weiß gekleidet, mit einer rosaroten Schleife am Busen* |
| | h - Moll | *schwermütige Weiblichkeit, die Spleen und Dünste brütet* |
| | d - Moll | *Ton der Geduld, der stillen Erwartung seines Schicksals und der Ergebung in die göttliche Fügung* |
| 2. Klasse | fis - Moll | *finsterer Ton* |
| (bedrückende Traurigkeit) | cis - Moll | *Bußklage, Seufzer der unbefriedigten Freundschaft und Liebe* |
| | gis - Moll | *Griesgram, gepreßtes Herz bis zum Ersticken; Jammerklage, die im Doppelkreuz hinseufzt; schwerer Kampf, mit einem Wort, alles was mühsam durchdringt ist dieses Tons Farbe* |
| | es - Moll | *Empfindungen der Bangigkeit als aller tiefsten Seelendrangs; der hinbrütenden Verzweiflung, der schwärzesten Schwermut, der düstersten Seelenverfassung* |
| 3. Klasse | g - Moll | *Mißvergnügen, Unbehaglichkeit, Groll, Unlust* |
| (hoffnungslose Traurigkeit) | c - Moll | *Klage der unglücklichen Liebe* |
| | f - Moll | *tiefe Schwermut, Leichenklage, Jammergeächz und grabverlangende Sehnsucht* |
| | b - Moll | *Moquerien gegen Gott und die Welt, Vorbereitung zum Selbstmord* |

Figure A1. Characteristics of musical keys in minor according to Schubart (1806).

**LIST OF FIGURE CAPTIONS**

Figure 1: Concept of musical consonance according to Terhardt (1984).

Figure 2: Interaction of spectral components in musical notation according to Helmholtz (1863).

Figure 3: Roughness of different musical intervals according to Helmholtz (1863).

Figure 4: Concept of sensory pleasantness according to Terhardt and Stoll (1981).

Figure 5: Illustration of the concept of "basse fondamentale" of Rameau (1750)

Figure 6: Illustration of the concept of virtual pitch of Terhardt (1974).

Figure 7. Characteristics of musical keys in minor according to Schubart (1806). English version of the original German given in Appendix A1. The (?) indicates difficulty in translating the exact sentiment. (Many thanks to Aaron Hastings, Rajani Ippili and Eckhard Groll, Purdue University, for this translation.)

Figure 8: Musical dynamics and category scales.

Figure 9: Loudness patterns for the note $f_4$ played by a horn in a reverberant chamber at the musical dynamics pianissimo pp (a) or fortissimo ff (b).

Figure 10: Loudness pattern of a piccolo flute (a) and a pipe organ (b) at same A-frequency weighted level of 85 dB(A).

Figure 11: Schematic plan of the ground floor of a concert hall with indications of the source S and three positions 1 through 3 of the receiver.

Figure 12: Rating of speech quality in a concert hall at positions 1, 2, and 3 for acoustic presentation alone (unfilled symbols) or acoustic plus visual presentation (filled symbols).