SOUND SOURCES

Orchestras and chamber music ensembles
Musical instrument acoustics
Musical instrument dynamic range and loudness
String instruments
Woodwind instruments
Brass wind instruments
Percussion instruments and piano
The speaking and singing voice

For the realization of proper concepts in recording, it is a basic requirement that one have knowledge of the acoustical properties of sound sources, especially those of the human voice, of musical instruments and of instrumental ensembles. These properties may be investigated initially without consideration of the sound radiation. However, it is just those radiating characteristics of the sound sources in a recording that provide important modifications of the sound properties. The radiating characteristic significance in wind instruments is greater than in string instruments, since radiation from instrument bells or holes provide a more pronounced radiating behavior. The radiating characteristics in the near field of an instrument are so significant that they have a greater influence on the total sound picture than does the choice of a microphone. The simplest way of selecting the optimum microphone is to place two microphones at different locations and to compare sounds by switching between them. The greater importance of proper microphone location over proper microphone type is most applicable to single microphone recordings and to recordings using support microphones. In recordings made at a greater microphone distance, such as is usual in serious music recording, the influences of the recording space diminish the significance of the sound source's radiating characteristics, so that one can then pay closer attention to the selection of a suitable microphone. It is not possible to recommend specific microphones for specific applications within the purviews of this text, but general information will provide some guidance in this direction.

Orchestras and chamber music ensembles

Instrumental configurations

If more than about 10 instruments form an ensemble, one generally speaks of an orchestra. The string voices usually account for multiple seats in such an ensemble. The membership of the symphony orchestra as a larger, and the chamber orchestra as a smaller, group results from the selection of the composition to be performed. The size of ensemble generally remains within a range of members indicated by the style and period of the composition but leaves certain details open. The core of the orchestra is the string sections, comprising Violin I, Violin II, Viola, Violoncello (or cello) and Bass. The cello and bass play the same parts for music composed prior to 1800, after which time separate parts were written for each of them. This string arrangement has remained unchanged till today. What has changed is the size of each section. As the wind sections were increased during the 19th century, the string sections had to be enlarged as well to maintain proper tonal balance (fig.A). During the baroque period, the ensemble was ofttimes enlarged through the addition of 2 oboes and/or flutes and 1 bassoon, sometimes by 2 to 3 trumpets and a pair of timpani depending on the style of the music. Baroque music always has the so-called *figured bass*, consisting of one or two bass instruments (cello, double bass, bassoon) and an instrument capable of playing chords (harpsichord, organ, lute). After 1750 the wind sections of the orchestra show steady growth, until around 1900 the orchestra reaches a size which can no longer be exceeded (fig.A). The double woodwind section in classical music, aside from some other expansions, has remained standard. It is only the percussion group which has been expanded in the 20th century.

All serious music groups of no more than 10 soloist voices are called chamber music ensembles. Instrumental music from before 1600 must be considered chamber music. We find here neither standards for participating musicians nor other clear instructions for the playing of compositions. Prior to 1600 it is the wind instruments which predominate and which were available in much greater variety, but with considerable playing difficulties. Chamber music has developed certain group memberships for baroque, classical and romantic music:

Solo sonata: composition for an unaccompanied instrument.

Duo, duo sonata: composition for two instruments. In compositions for piano and another instrument, the piano is usually not mentioned: a violin sonata may therefore be either a sonata for violin and piano, or one for unaccompanied violin. A duet is a composition for two melody voices.

Trio: composition for 3 instruments or voices. A frequent form is the piano trio (violin, piano, cello); a horn trio may be a work for three horns or for one horn and two other instruments. The trio sonata in baroque music was composed for four musicians (2 melody, 1 bass, and 1 chord instrument). (Editor: The German word *Terzet* refers to a composition for three singing voices.)

Quartet: composition for 4 instruments. One of the most important such groups is the string quartet (violin I, violin II, viola, violoncello).

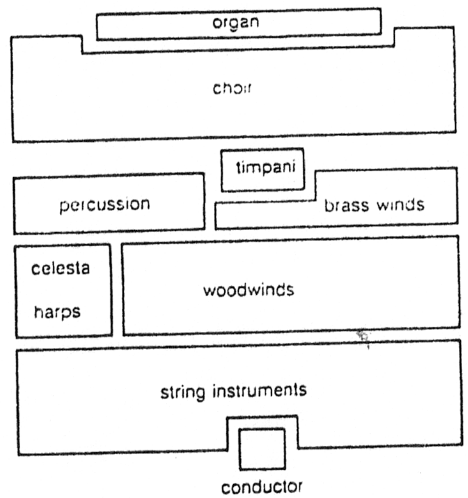
Quintet, sextet, septet, octet, nonet, and decet: compositions for 5, 6, 7, 8, 9, and 10 instruments.

In the area of rock, jazz and popular music it is difficult to speak of the instrumental configuration in studio productions, because the number of voices or instruments often does not agree with the number of musicians.

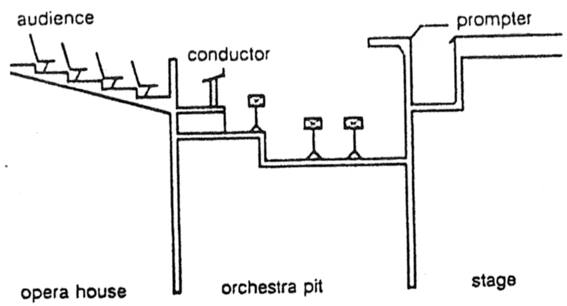
A. Standard symphony orchestra seating development (strings number only approximate)

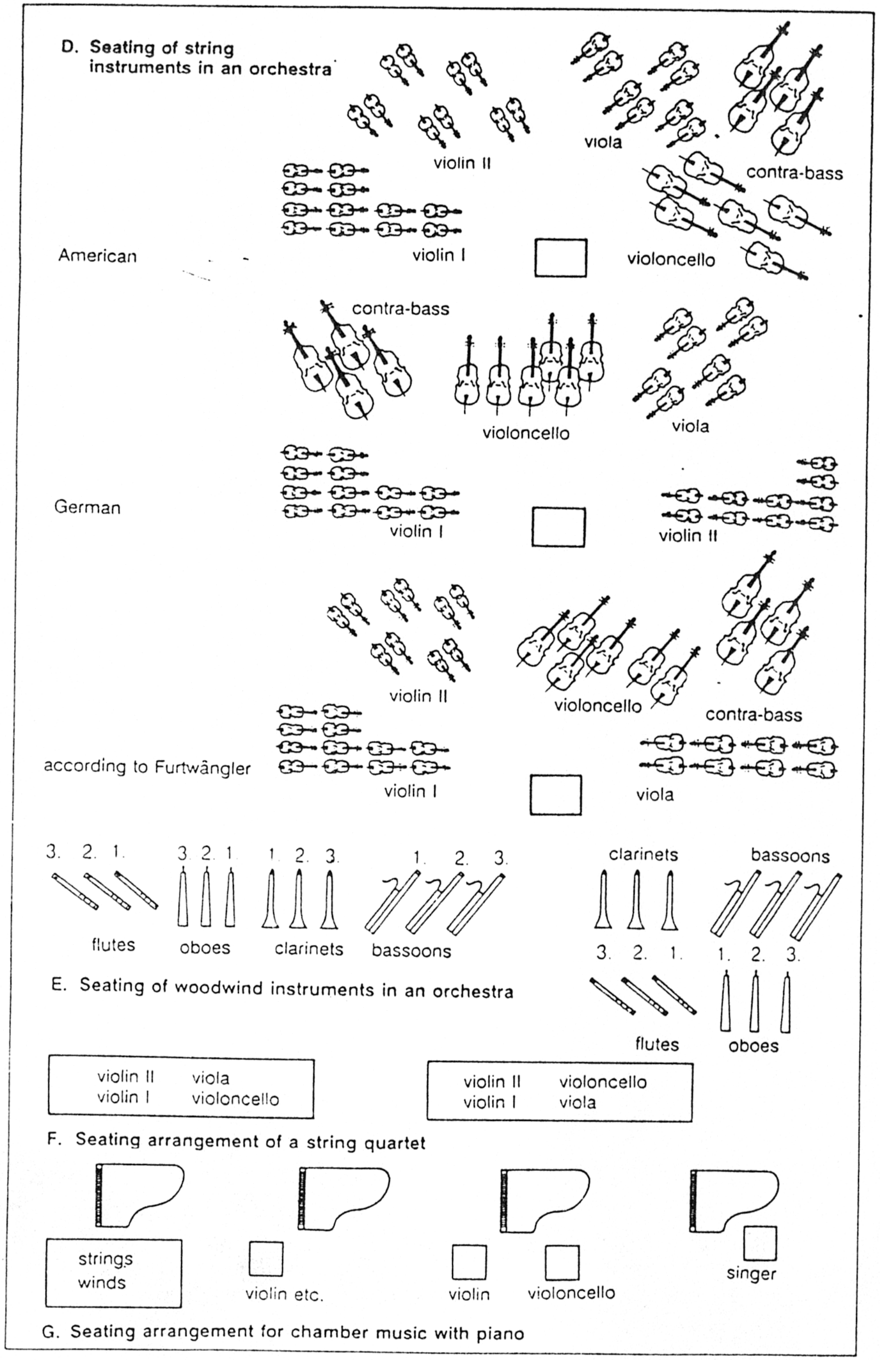
Instruments	Woodwinds								Brass				Strings				
Period	flutes	oboes	clarinets	bassoons	piccolo	english horn	bass clarinet	contra-bassoon	horns	trumpets	trombones	tuba	violin I	violin II	viola	violoncello	contra-bass
1750-1770		2							2				8	8	6	4	2
1770-1790	1	2		2					2								
since 1790	2	2	2	2					2	2	3		12	10	8	7	6
since 1840	2	2	2	2	1	1	1	1	4	3	3	1					
since 1870	3	3	3	3	1	1	1	1	4	3	3	1	16	14	12	10	8

B. Seating order of a large symphony orchestra with chorus



C. Opera house orchestra pit





This is due to the so-called *playback* recording process. There are vastly differing numbers of instruments depending on the style and time of the composition, the arrangement or the production. Some sort of standard has developed only with a few musical styles. Basically the requirement is for a more or less elaborate percussion group, a bass instrument (electronic, string), and an instrument capable of producing chords (guitar, keyboard). In the most recent decades it is the electric guitar which has largely determined the sound; often as many as 3 electric guitars are used. Wind instruments also play an important role (trumpets, trombones, saxophones, clarinets). In addition such groups also include acoustical or electronic keyboard instruments, while strings rather form the exception. The big jazz band also has no standard group configuration, but one can use the following as a guideline: 4 trumpets, 4 trombones, 5 saxophones, 1 clarinet, 1 guitar, piano, bass and drums. Brass bands also have differing configurations; of first order of importance we find brass (trumpets, trombones, horns, tubas, alto horns, tenor horns, baritone horns), often augmented by woodwinds (piccolo, flute, clarinet, bassoon). Add to that bass and snare drums, cymbals, etc.

Seating arrangements

Fig.B shows the stage arrangement for the different instrumental groups of the symphony orchestra, and fig.D and E the seating within the instrumental sections. Since the 18th century and until 1945, Germany used the German or classical seating, while after 1945 the American seating, basically introduced by Leopold Stokowski and also suggested in a variation by Wilhelm Furtwängler, became popular. The American seating is particularly dedicated to precision playing, while the German seating appears to have advantages in spacial balance, particularly important in stereo recording. In the opera orchestra the limited space of the orchestra pit determines the seating (fig.C), which is not as standardized as is the one on stage. Figs.F and G show examples of the seating of chamber music groups. For rock or popular music the acoustics of the studio determines the seating (→p.94), while on stage it is mostly dictated by the visual effects required in show business.

Dynamic range and level

The dynamic range of a large symphony orchestra depends on the dynamics of the individual instruments on the one hand, and on the composition and the acoustical properties of the reproducing space on the other. For live concerts with audience, the background room noise has a direct effect on the softest possible sound pressure level of the music in the hall: about 35 to 45 dB and only slightly above the room tone. The highest sound pressure levels are generated by the brass and percussion, lying seldom above 100 to 110 dB. Thus the dynamic range of a large orchestra may be between 50 and 75 dB. For works of the 18th century and the commensurately smaller orchestra, the dynamic range is smaller, about 35-50 dB. Studio recordings allow a greater dynamic range, since the extreme quiet in the studio permits unmasked lower sound levels. Studio recordings with over 80 dB dynamics are routinely possible.

The total output level of an orchestra, as perceived by the audience, is dependent on the number of players, as well as on the particular room, since the diffuse sound level increases directly as a function of the reverberation time but decreases with increasing room volume. For a recording, on the other hand, the total level is principally a function of the number of players, since the direct sound predominates. The level of short duration tones, when compared to that of longer duration tones, is largely room dependent as far as the concert goer is concerned, but is of no practical significance in a recording.

Time analysis of a tone

A musical tone is perceived as an organic whole, but acoustically it is composed of three segments with differing structures which follow one another sequentially: attack, quasi-stationary condition, and decay (fig.A). While all instrumental notes naturally begin with the attack, it is only the string and wind instruments which have a quasi-stationary condition, during which time their acoustical behavior changes but little. All percussive instruments, to which belong the piano and all plucked instruments such as the harpsichord, transition directly from the attack into the decay phase. An exception to this is the electric guitar: it can die away so slowly as to give the impression of a quasi-stationary phase. Wind and string instruments have but a relatively short decay time. The three-part time behavior of a tone, therefore, corresponds to the reaction of a space to a sound event, with reverberation build-up (attack), quasi-steady state, and decay. The shaping of the sound resulting from the resonant body of the instrument is modified analogously through influences of the surrounding room.

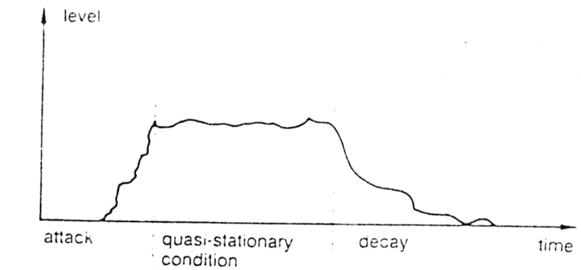
The attack is a significant portion of a tone; it contributes to the recognizability of instruments and lasts between 1 and 250 ms, depending on the sound generator and the attacking system (fig.B). Under 10 ms it has the characteristic of a click which appears at relatively high level quite separate from the tone itself. Between a duration of 10 and about 40 ms, the attack cannot be perceived at all; above that it sounds soft and the build-up of the oscillation becomes clearly audible. The shorter the attack, the noisier it is. The attack which results from the diagonal editing cut of a 38.1 cm/s (15 ips), ¼" tape creates an artificial attack of 17 ms and is therefore in the unnoticeable range. Since such edits generally are chosen to come just ahead of a loud section, the backward masking effect can also serve to cover it.

With decay the level is reduced exponentially as a function of the specific instrument, just as is the case with reverberation, while the sound spectrum changes at the same time. The higher harmonics decay faster, making the sound duller with increasing decay time (fig.C), something which is true for the room reverberation as well (→p.22). It is easy to see why we use the terminology of reverberation, in view of the similarity between tonal and reverberation decay. For those stringed keyboard instruments which are of particular interest because of their decay characteristics, we obtain decay times of up to 40 seconds at low frequencies, while they reduce to only a few seconds at higher frequencies. Therefore, the instrument's own decay time is quite long when compared to the reverberation decay time. This means that the echoiness of rooms will not be as distinct for these instruments as it is for wind and string instruments.

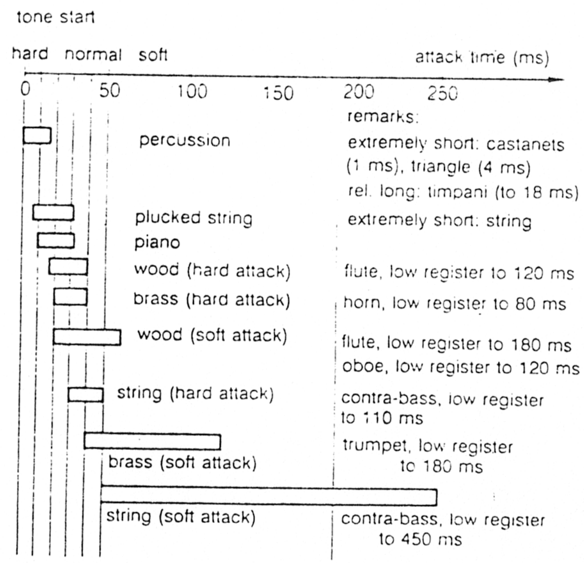
Quasi steady-state oscillation characteristics

The quasi steady-state for instruments, with the exception of electronic ones, is never an even one, but rather displays typically regular and also especially irregular unsteadinesses of the most important attributes of sound. The tremolo refers to a strong amplitude modulation of the tone, which, for string instruments, is produced by a rapid oscillation of the bow against the string and in wind instruments by the so-called flutter-tongue. In its pure form the vibrato in actuality is a frequency modulation but in musical instruments it is almost always combined with some amplitude modulation. The vibrato is used by all string and wind instruments (with the exception of the clarinet and french horn) to improve the tonal quality.

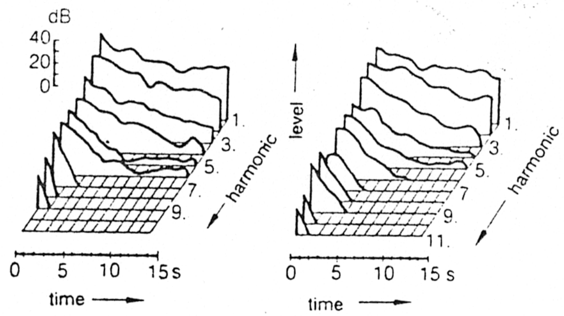
A. Sections of a musical tone

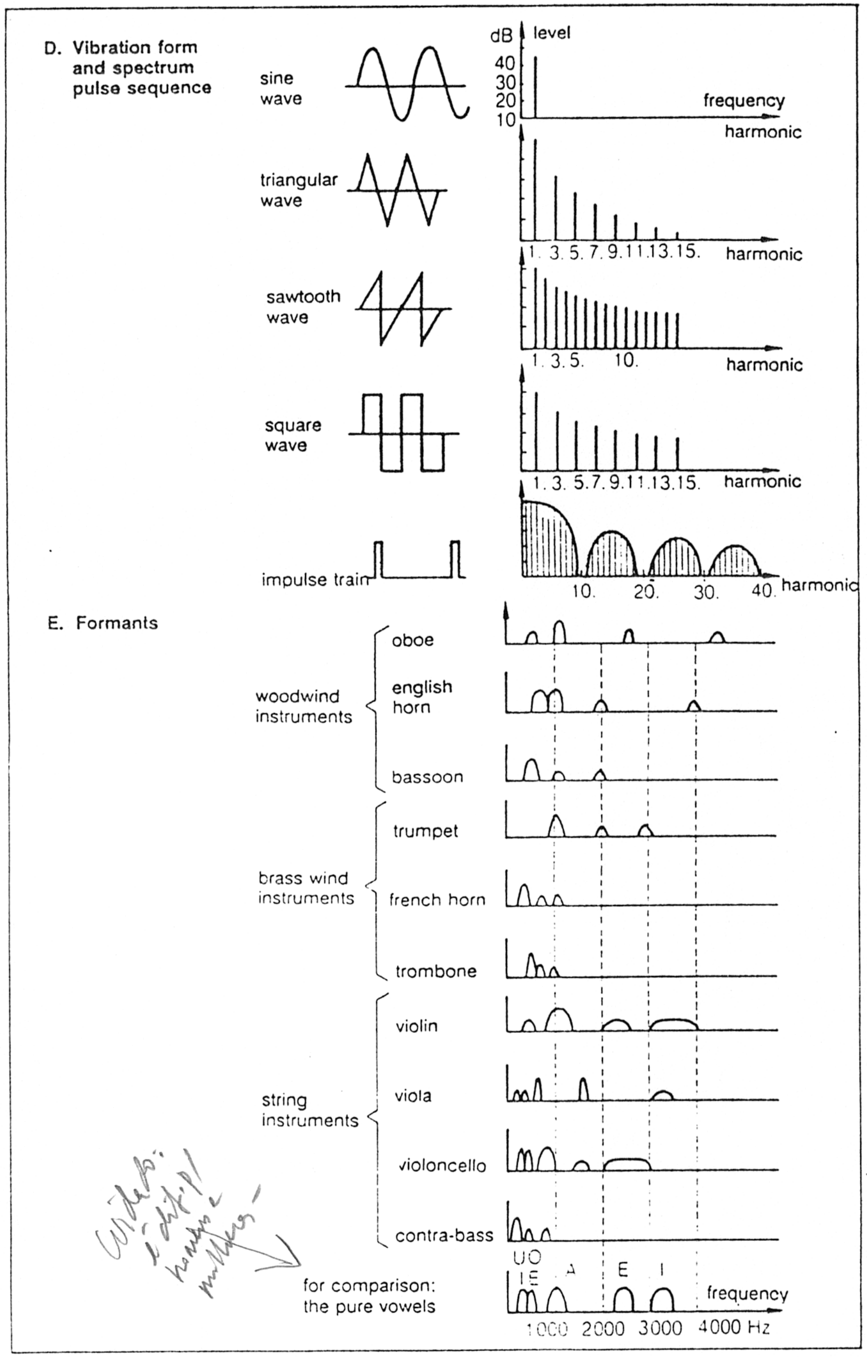


B. Attack time of musical instruments



C. Decay of musical instruments (piano)





All wind instruments including the organ, and all string instruments generate musical tones (sounds) which are made up of so-called harmonics (partials). These partials have a harmonic relationship, i.e., the frequencies of the harmonics are a whole integer multiple of the frequency of the first harmonic or fundamental. Deviations from the regularity of this spectrum are found particularly in the piano sound. The reason is the thickness and stiffness of the strings. Oscillating plates or pipes such as cymbals, tom-tom, gong or bells, and oscillating membranes such as kettle drums (timpani) and all types of drums, but also rods such as triangle, display a more or less non-harmonic spectrum. They mask the sensation of tones as in bells, gong, triangle or timpani, or give the sound an entirely noise-like effect as with cymbals, tom-tom and in drums. The frequency range of the spectrum not only depends on the type of instrument and on the location of the microphone or listener, but also on its placement within the dynamic range (→p.40). The frequency range may increase by a factor of three to ten in the transition from pianissimo to fortissimo. The lowest register instruments, with fundamentals starting at about 25 Hz are the double bass, contra-bassoon, bass tuba, organ, bass drum etc. The usual bass instruments range begins at about 60 Hz. The upper boundary of the spectrum for double bass, contra-bassoon, bass tuba and timpani is at about 5000 Hz; for most other instruments between 10 and 15 kHz while some percussion instruments such as the triangle go beyond that (→p.31).

The wave form – fig.D shows some idealized basic wave forms – determines the intensity and frequency range of the harmonics, i.e., the spectrum envelope. It also determines whether all harmonics or only the odd ones are present. The idealized wave forms shown in fig.D only exist in electronic instruments, but natural wave forms display enough similarities to permit a derivation of the natural wave forms. The flute comes closest to the sine wave form, especially when playing softly, while the string instruments follow more the sawtooth wave form. Clarinets are most similar to square wave forms while double reeds and, to a minor extent, the brass instruments display a wave form closer to pulses.

Formants are resonance-like, amplified harmonics which have a fixed position in the spectrum regardless of the tone being played. Formants are particularly an attribute of string instruments, double reeds and brass (fig.E). They are also responsible for the differences in speech vowels (fig.E). It is for this reason that the formants impart a kind of *vowely* quality to the instrument sound. The bright, open vowel sound "ah" is to be found in the violin, trumpet and oboe. The bassoon sound is best described by the "oh" formant. The so-called nasal formant is located in the range of 1800 to 2000 Hz and is highly developed in the saxophone.

The background noise is as much a part of the sound as are the harmonics. It is strongest in string instruments and displays the resonant quality of the particular instrument. Somewhat less, but nevertheless quite indicative, is the noise component of the woodwinds. The flute is accompanied by a rather typical breathiness while the noise components are weakest in the brass. These noise components are found at a level of about 30 to 50 dB below the level of the strongest harmonics. The noise modulates the amplitude behavior of the sound which thereby undulates by several dB. A sound's noise components are much more important in recording than for the live listener since microphones are much closer to the instruments than listeners are. They increase the instruments' presence and intensify the sound. Breathiness in wind instruments identifies an individual player.

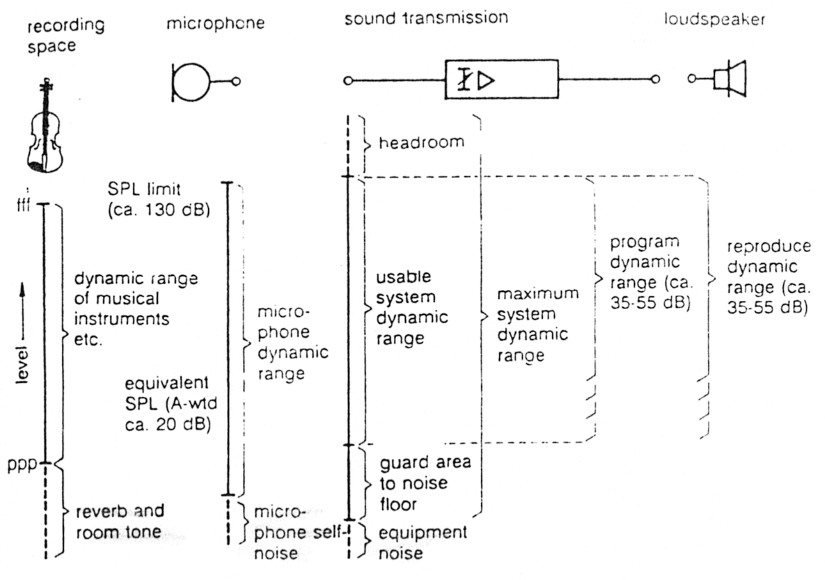
Musical instrument dynamic range and loudness

Technical dynamics

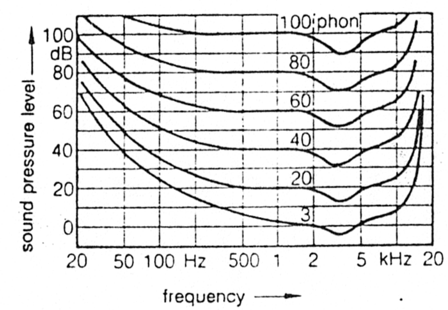
In electro-acoustics the dynamic range is generally defined as the level range between an upper and lower boundary level. Analogously we have the dynamic range of musical instruments and similar sound sources as the difference between the highest and the lowest producible sound levels. When the room tone is included, something one can hardly separate from the instrument sound in practice, the lower boundary level is hardly definable because the sound level, together with the reverberation decay, dips into the unavoidable background noise. The technical dynamic range is firstly made up of the microphone dynamic range which lies between the overload point and the self noise level. System dynamic range describes the dynamic values of the entire technical transmission chain. One must differentiate between the maximum and the effective system dynamic range which describes the true usable level range considering the overload reserve and a reasonable number of dB from the base noise. One of the problems in quoting the technical dynamic range is the fact that there are several standard methods of noise measurement which may yield differences of up to 10 dB. When measuring the highest usable level in analog transmission or storage, it is the peak program level meter (PPM) displaying *quasi-peak* levels which is the most suitable instrument, while for digital transmission or storage one should use a *real-peak* level indication which may yield up to 10 dB higher peak levels. When using the usual PPM for digitally processed signals, one must allow for an overload reserve. The program dynamic range is the dynamic range which is established, considering the desired or possible playback dynamic range (fig.A.)

Level and loudness

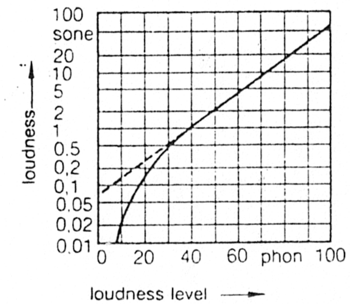
The levels which define the dynamic range are physical magnitudes which are important to the audio field. However, during the enjoyment of music and speech the physical magnitudes are subjectively evaluated as well by utilizing our prior audio experience. The values of dynamics, therefore, only provides a first, rough estimate about the possible differences between sensed loudnesses. Loudness as a measure for the subjectively sensed sound intensity is only defined for sine waves. This is done using contours of equal loudness and finding the sound pressure level of the equal loudness sine wave tone at 1000 Hz. This is also the loudness level in phon. By means of the loudness curve, this value is then converted to loudness in sones (fig.C). Non sine wave signals may be converted to loudness values by loudness comparisons with sine wave signals or through application of certain standardized calculations. In general, a level increase of about 10 dB corresponds to a doubling of the loudness; a doubling of the sound pressure or the voltage leads to an increase of only 6 dB. Only a tripling of the sound pressure level or a multiplication of the sound power by ten doubles the loudness. An increase in the number of instruments playing at equal volume, for example violins, leads to the following level increases: twice the number of players provides 3 dB more; four times the number provides 6 dB; eight times 9 dB, ten times 10 dB, and sixteen times results in a rise of 12 dB in level. This shows that for a significant increase in loudness, one must provide an inordinately greater number of musicians. The level relationships are exactly reversed when several microphones with about equal output level are to be mixed together. If two such microphones are mixed then the level of each one must be 3 dB below the desired total level, for 4 microphones it would be 6 dB and so forth.



A. Dynamic range, concepts



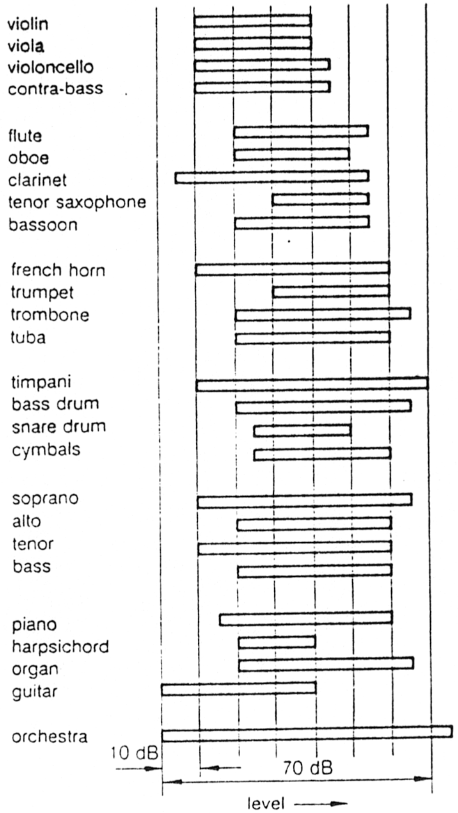
B. Equal loudness contours for frontally impinging sound



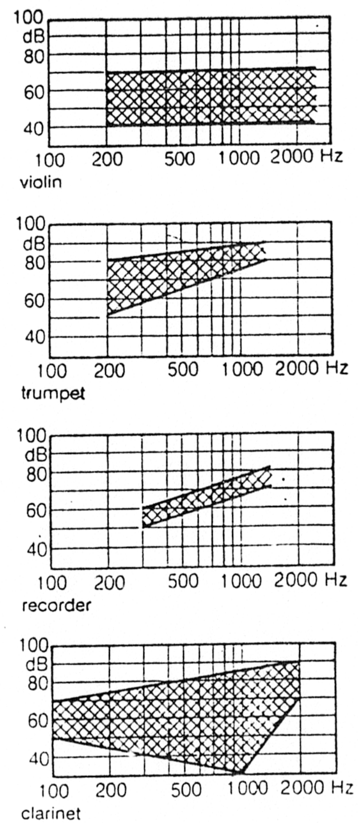
C. Relationship between loudness frontally impinging sound

soft → loud		
ppp	pianissimo possibile	as soft as possible
pp	pianissimo	very soft
p	piano	soft
mp	mezzo piano	half soft } identical
mf	mezzo forte	half loud } identical
f	forte	loud
ff	fortissimo	very loud
fff	forte fortissimo	as loud as possible

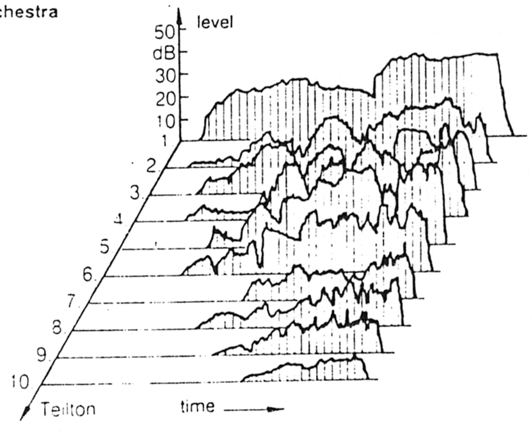
D. Dynamic steps in music



E. The dynamic range of instruments, voice and orchestra



F. The pitch dependence of dynamic range



G. The transition from pianissimo to fortissimo of a tone of increasing loudness (clarinet)

Two further factors influence the relationship between level and loudness: (1) for sound components below 500 and above 5000 Hz the ear is less sensitive; and (2) for program material with like meter indications, sound level patterns with numerous spikes appear to be softer than those without spikes.

Level and timbre dynamics

The dynamic range of a musical instrument or ensemble is categorized in steps of musical dynamics. They reach from the lowest playable to the highest playable loudness (fig.D). The individual dynamic steps are differentiated in their loudness or acoustical level (level dynamics) and in their sound coloration or spectrum (spectrum dynamics). The level dynamics describes only the level differences. Between two steps of the dynamic range there usually exists a 6-10 dB level difference. The absolute sound level differences in live music are dependent on several factors: on the sound level of the sound source (a function of the particular instrument and playing technique), on the distance, on the reverberation time, and on the room size. As a result, the sound level values may vary within wide limits. Of critical importance, especially for the listener (less for the recording) are the sound levels of those instruments played in large rooms which are too weak, since their low frequency sound components sound even lower due to the sensitivity losses of human hearing. Fig.E shows the dynamic range of musical instruments, the dynamic range of the human voice, and that of a large symphony orchestra under actual conditions. On average, the string instrument level is about 10 dB below the level of the wood winds while these are 10 dB below the level of the brass. For an orchestra this is equalized through commensurate membership of the various instrument sections to achieve a balanced sound picture. For many instruments the dynamic range is a function of the register in which they play. Fig.F shows typical dependencies. For strings, but also for piano, guitar and harp, the dynamics and absolute levels are fairly close over the entire range. For the brass the dynamic range in the upper register becomes smaller while the absolute level increases. With flutes, and especially recorders, the dynamic range is evenly small while the level increases with increasing frequency. The clarinet, by comparison, has a rather large dynamic range in its mid-range. These acoustical instrument phenomena must be considered by the experienced composer or arranger.

The timbre or spectral dynamics are just as important in music and recording techniques as are the level dynamics. Because of the fact that every dynamic step is characterized by a specific spectrum, it may be recognized independently of the audible level. This is a basic prerequisite for the performance of music altogether because a forte passage at a close listening distance will never become piano at a greater distance. The number and intensity of the harmonics increases with loudness. The formants, too, are built up one after the other. This loudness defined by the spectrum is therefore in reality not a true loudness, but rather information about how loudly the instrument was played. It is this very spectral information which makes it possible to alter the level for technical reasons altogether, i.e., to reduce the dynamic range. Going too far with such dynamics restriction causes contradictions between spectral and level dynamics which are not acceptable. Fig.G gives an example of how the spectrum of a swelling tone changes for a transition from piano to forte.

The ear normally judges loudness by utilizing the distances of sound sources. Therefore it does not determine whether two sound sources are equally loud to the listener at his listening position but whether they are truly equally loud. The ability to judge loudness, of course, is meaningful for our orientation in the world around us. It also leads to the fact that the sound pressures at the ear when listening with headphones is considerably higher than that from loudspeakers for a sensation of equal loudness.

*

Strings

Instruments

There are four prevalent string instruments: violin, viola, violoncello (or cello), and contra bass (or double bass). The violin, viola and cello differ mainly through their size but only minimally through their shape. The double bass is built differently in several details. Besides these, we occasionally use historical string instruments such as the viol or gamba which is played on or between the knees, and from this group of instruments especially the tenor gamba which is most like the cello in its pitch range and size.

Application

In a symphony or chamber orchestra, the string instruments are choral, meaning that there are several instruments in each section. There are traditionally five voices: violin I, violin II, viola, cello and bass. The membership of these sections in large orchestras and in the order given is 16-24, 14-20, 12-16, 10-14 and 8-10, and in small orchestras 8-10, 6-8, 4-6, 4-6, and 3-4. These groups of string players form the core of the orchestra. In chamber music they form the most popular instruments along with the piano, but here each instrument (except for the violin) usually appears only once. There are relatively few compositions for a single, unaccompanied string instrument. In jazz the normally plucked bass is of great importance, and recent folk music influences have also returned the violin to this art form. In big bands, the strings serve about the same function as in a classical orchestra. And in popular music strings are used for background effects or as soloists; however, they serve a subordinate role here and, except for their use for solos, are often replaced by electronic strings which either synthesize the string sound or have memory circuitry containing the sounds of real strings.

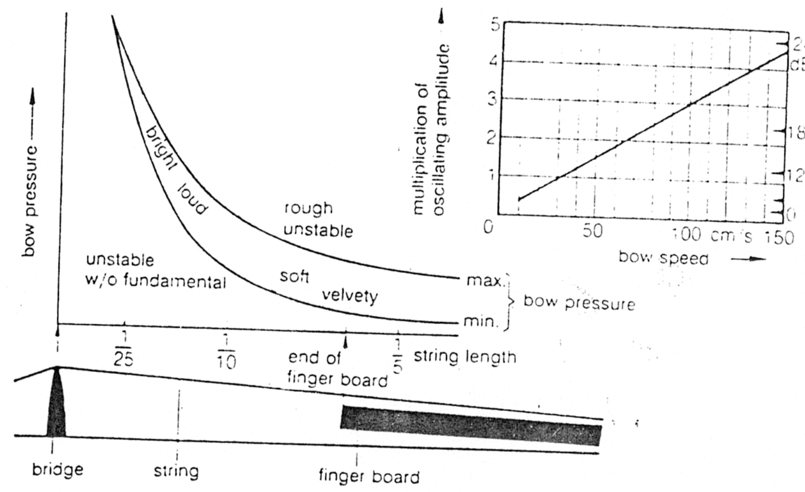
Sound acoustics

Sound generation: Because the string sticks to the rosined hairs of the bow and then periodically rebounds when its excursion becomes too large, we obtain sawtooth like string vibrations. These contain the total sequence of harmonic partials at high amplitude as their acoustic raw material which abates with increasing frequency (–p.36). The vibrations are transmitted through the bridge to the resonance body which, in turn, radiates them to the environment. This resonance box is part of a very complex resonant system which greatly modifies the original form of vibration. To shape timbre and loudness, the player may change the speed, pressure and position of the bow. The intensity is really only influenced by the speed of the bow while its pressure and position on the string determine the timbre (fig.A).

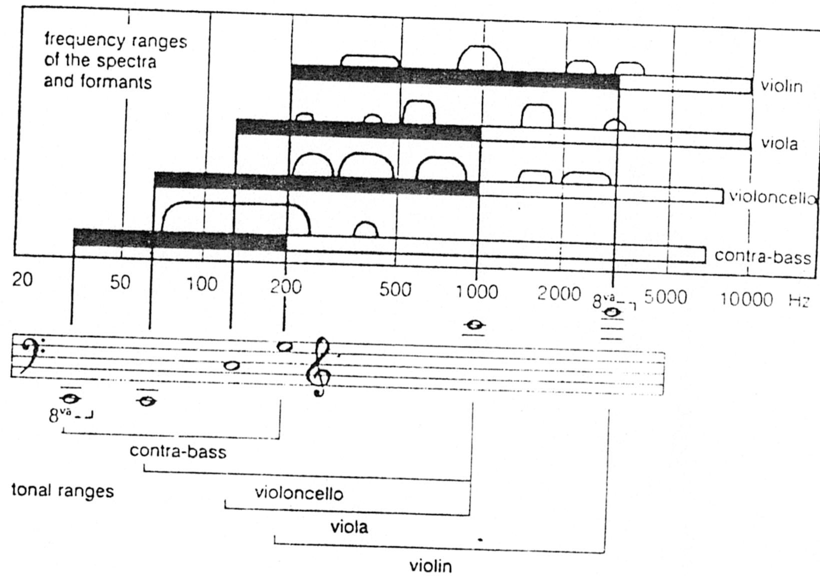
Attack: Among all of the instruments the strings have the longest duration attack phase. For normal playing, the tone is fully formed only after 100 ms (for the double bass after 400 ms). For a sharp down beat, these times are reduced to 30-60 ms and 150 ms respectively. This may explain why one often hears the comment that the bass voice appears somewhat delayed. Pizzicato tones have a much shorter attack of under 20 ms and therefore give the impression of greater precision than bowed tones.

Tonal range (fig.B): The tonal range is limited at the bottom by the tuning of the lowest string while the upper boundary is strictly a function of the musician's skill.

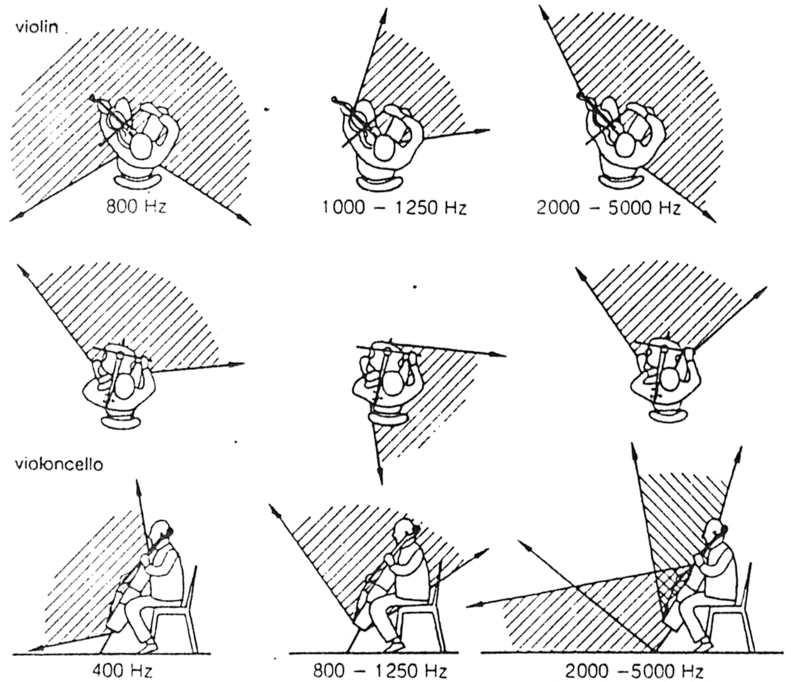
Frequency range (fig.B): In string instruments the sound partials beyond 10 kHz are relatively weak. The individual frequency range is highly dependent on the playing method. The response range widens towards the high end with increasing bow pressure



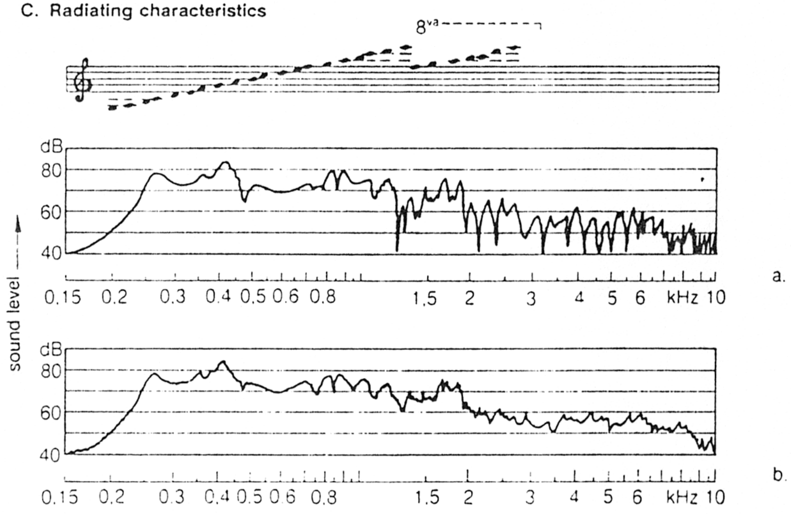
A. The influence of playing technique on the sound of string instruments



B. Frequency ranges of the spectra with formant positions and tonal ranges



C. Radiating characteristics



D. Frequency response of a violin
a. in the near field at a particular microphone location
b. In the near field for super imposition of several frequency response curves of various microphones.

and decreasing distance of the playing location to the bridge (fig.A).

Formants (fig.B): The resonant properties of the sound box, its form, dimensions, materials, and its construction determine the frequency position of the formants, i.e., the areas of accentuated sound components. The resonance areas are fixed while the frequencies of the played tones change constantly, so that constantly changing partials fall within the range of the resonances. As a result, the timbre may vary greatly from tone to tone. Excellent sounding instruments have no sharp resonances and therefore have a unified character for all the tones. It is the formant in the vicinity of 1000 Hz which gives the violin its characteristic sound. It provides the bright, open characteristic which is found also in the vowel "ah" and which makes it the preferred vowel for singing (la, la, la). The nasal formant between 1500 and 2000 Hz gives the viola its characteristic sound. The cello also has a certain nasal character but is most significantly described by the width of its formant range of between 2000-3000 Hz, which lends the instrument a certain sharpness which allows the cello to sound brighter at higher frequencies than the violin in spite of its larger size.

Noise components: Typical is the relatively strong attack noise which, for wind instruments, may be 20 to 30 dB louder than the wind tones (with the exception of the flute). This noise has a continuous spectrum which mirrors the resonant properties of the instrument identically with every tone. Of special note is the buzzing sound of the double bass which lends it its special sound within the orchestra and which is caused by the vibration of the bow hairs. The noise components of the string instruments are largely independent of the playing intensity, i.e., for softly played tones they are relatively the loudest.

Dynamic range and level: The dynamic range is relatively even over the entire tonal range (→p.40). It only gets a bit restrictive towards the upper end of the frequency spectrum. The sound levels are quite low when compared to the wind instruments; on average 10 dB lower than the woodwinds and 20 dB lower than the brass. These differences are compensated by choice of the number of players in each section (→p.40).

Radiating characteristics

The radiating characteristics are basically caused by vibrating wooden sections of the resonance or sound box at various amplitudes and in various phase relationships. This results in relatively complex behavior which varies within certain limits (fig.C). The radiating pattern in the frequency area of the formants is restricted to a small angle while at low frequencies we find an omni-directional radiating pattern. Towards the upper end of the response range, the radiating pattern is not narrow like the brass instruments' if one examines a wider response range. The frequency response for a microphone position at close range shows something of a comb filter in its micro-structure which may lead to an unnatural edginess. It is only the superimposition of the response curves from all directions – something best accomplished by reverberation – which smooths out the frequency response curve (fig.D). Therefore, it is recommended that microphones be placed at a greater distance from strings, especially for serious music. The basic difference between natural and artificial reverberation is especially pronounced in string instruments: while the natural reverberation represents an integration of all the radiating directions of the instrument, the artificial reverberation imparts the specific frequency response at the location of the microphone.

Woodwinds

Instruments

The woodwind instrument section encompasses a relatively large number of different instruments which differ greatly in the way they produce their tones. Not all of them are made of wood. The flute is made of a silver amalgam and the saxophone was made of metal right from the start. The most important woodwinds today are the flute, the oboe, the clarinet (in Bb, and more rarely in C or A), the bassoon, the english horn and the saxophone. Add to that those instruments derived from the main ones: from the flute family the piccolo and alto flute; from the oboe group the oboe d'amore; from the clarinet family the small clarinet, the basset horn and the bass clarinet, the contra bassoon, and from the saxophone family the alto, tenor, soprano, baritone and bass saxophones. Historical woodwind instruments for baroque music are the transverse flute which is the predecessor of the modern flute, the various recorders, and baroque oboes and bassoons. The instruments of the baroque era contain a multitude of recorders and other flutes, but especially double pipe instruments of different constructions and timbres.

Most woodwinds are written in a tonality different from the one which is heard. In other words they transpose.(fig.A).

Application

In the classical symphony orchestra we normally find two each of flutes, oboes, clarinets and bassoons. For music prior to 1800 their number is reduced. In the 19th century their number has gradually increased through addition of the piccolo flute and contra-bassoon, then the english horn and the bass clarinet (→p.32). Chamber music with exclusively wind membership stems solely from the 18th century. Strings were added after that. Generally speaking woodwinds are not as important as strings in chamber music. In pop and jazz the saxophone and clarinet have found popularity with the flute coming along as well. The rest of the woodwinds find only occasional use in popular music, but almost always in semi-classical music. Brass bands partly use woodwinds as well, especially clarinets, flutes, bassoons and saxophones.

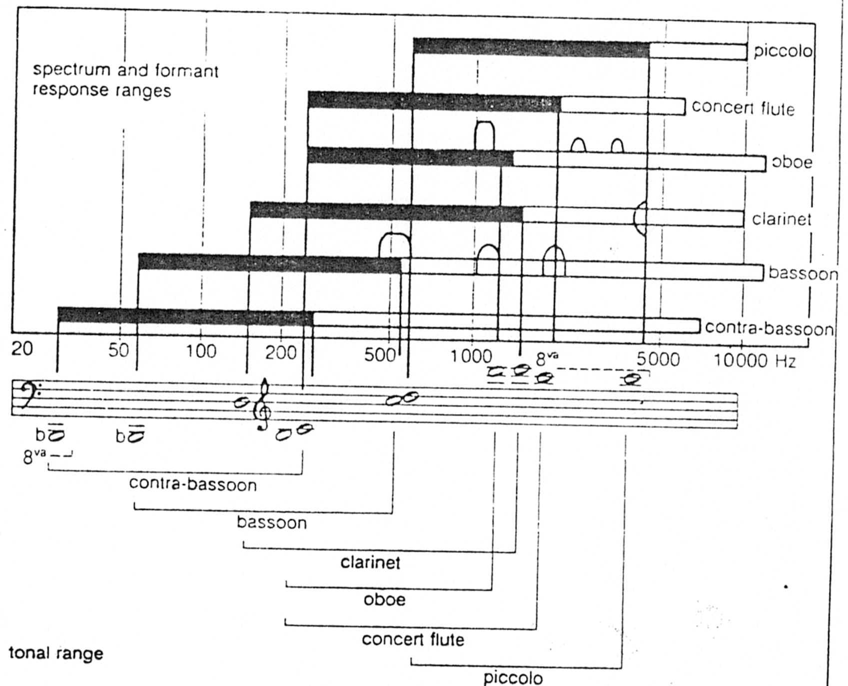
Sound acoustics

Sound generation: Woodwinds are divided into three groups according to their embouchure: the so-called vibrating air reed of the flute aims at a sharp edge and oscillates back and forth between the resonating air column within and the surrounding air without. As a result one hears a fairly constant embouchure noise which is typical for the flute sound. Clarinets and saxophones have a simple bamboo reed, which alternately opens and closes the instrument under the embouchure and lip pressure of the artist thereby exciting the air column in the resonant tube to oscillation. With oboes and bassoons this same function is performed by a double bamboo reed. The quality and suitability of the double reed is a constant problem for the oboe player – far more than with the single reed. Because of this, it is a common practice for oboists and bassoonists to make their own double reeds, i.e., varying the hardness or easiness in the embouchure considering the type and tonal range of the music to be played. Clarinet players, on the other hand, can make do with manufactured reeds.

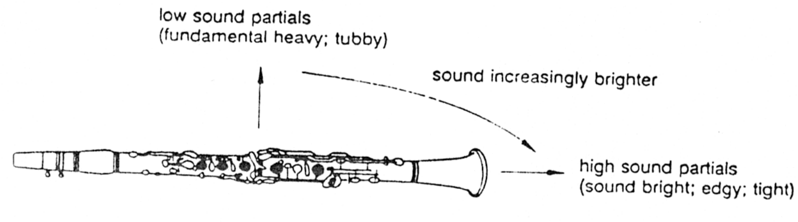
Attack: With the exception of the flute, the attack time of 10-40 ms is considerably shorter than that of the strings and, therefore, more precise. By contrast

Instrument		Notation	Pitch referred to notation
flutes	concert flute piccolo (small flute) alto flute in G (in F) soprano flute alto recorder	G-clef G-clef G-clef G-clef G-clef	as written one octave higher a fourth (fifth) lower one octave higher as written
oboes	oboe english horn oboe d'amore	G-clef G-clef G-clef	as written a fifth lower a minor third lower
clarinets	clarinet in B (in C; in A) small clarinet in Eb (in D) basset horn in F(in Eb) bass clarinet in B	G-clef G-clef G-clef bass clef	a major second lower (as written; a minor third lower) a minor third (major second) higher a fifth (major sixth) lower a major second lower
bassoons	bassoon contra-bassoon	bass clef bass clef	as written one octave lower
saxo phone	soprano sax in B alto sax in Eb tenor sax in B baritone sax in Eb bass sax in B	G-clef G-clef G-clef G-clef G-clef	a major second lower a major sixth lower a major ninth lower one cotave + major sixth lower two octaves + a major second lower

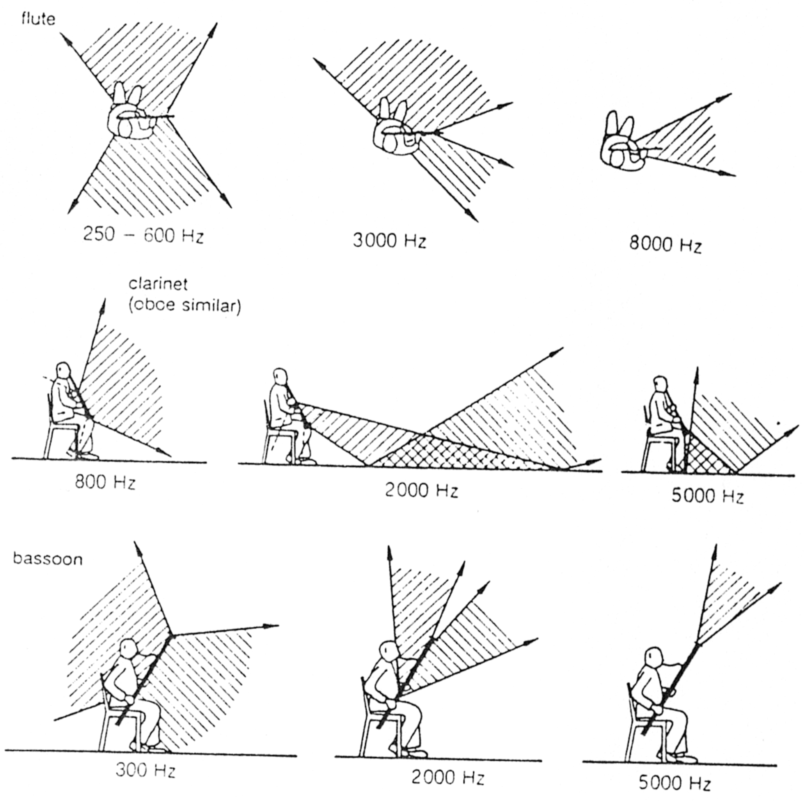
A. Woodwind instrument notations



B. Spectrum response ranges with formant positions and tonal ranges



C. Main radiating areas



D. Radiating characteristics

the flute, in its lower range, may display an attack time of over 150 ms and, therefore, its onset will appear commensurately soft.

Tonal range (fig.B): The tonal range encompasses two to three octaves, also dependent on the skill of the player. The contra-bassoon is the orchestra's lowest instrument and the piccolo its highest.

Frequency range (fig.A): The spectrum of the woodwinds reaches generally to 10,000 Hz. The flute has a distinctly smaller range of only about 6000 Hz. Of all the instruments of the orchestra it is the flute which has the strongest fundamental. It sounds *sine wavy*. The other instruments have strong 2nd and higher harmonics.

Formants (fig.B): Among the orchestral instruments, the double reeds are most strongly characterized by their formants which form during the embouchure due to the special form of oscillation of the double reed. The oboe has the same formants as the French nasal vowel sound "in". Its sound is therefore nasal and bright. For the bassoon the vowel sound "aw" is characteristic. The flute's weak formants give only little information about the difference between individual instruments, while in the clarinet they also have little meaning. The flute is largely recognizable by its embouchure noise and the relatively small amount of harmonics, the clarinet by its suppressed even order harmonics levels.

Noise components: The background noise is relatively weak when compared to the string instruments. The flute forms an exception with its typically flute-like embouchure noise.

Dynamic range and level: A strong pitch dependence of the dynamic range is typical for the woodwinds. Therefore even adjacent tones may have differing dynamic behavior. The flute displays only limited dynamics in its upper register. The oboe has this in the lower register. The clarinet has a very wide dynamic range in its mid-range. The level of the woodwinds is about 10 dB higher when compared to that of the strings, meaning that they sound about twice as loud. The levels increase somewhat with increasing pitch, especially for the flute (→p.40).

Radiating characteristic

Summarized simply, the lower and mid-range tones up to about 2000 Hz, emanate sideways from the finger holes of the instruments. Higher pitched ones beginning at about 3000 to 4000 Hz radiate from the bell (fig.C and D). The tonal timbre changes more severely with the radiating direction than is the case with the string instruments.

It is therefore more important to select the proper microphone position rather than the most suitable microphone type. This selection becomes even more decisive the closer the microphone gets to the instrument. With increasing instrument-to-microphone distance, the diffuse sound merges all of the radiating directions into a total sound, which becomes ever more independent of the microphone position. A close miking position absolutely requires that the instrument always be held in a fixed position, something of which normally only the best studio musicians are capable. *

The flute acts like an acoustical dipole since it also radiates sound from the mouthpiece. This causes certain sound cancellations within narrow angles, which, in turn, makes it necessary that a microphone be placed in such a position that it is equidistant from both ends of the instrument. By contrast to the other woodwinds, the flute emits audible breathiness from its mouthpiece.

In a saxophone the radiation from the finger holes and from the bell coincide because the bell, with the exception of the soprano sax, is directed upwards. This is also true for the bass clarinet.

Instruments and their application

The following brass instruments are found in a symphony orchestra depending on the style and era of the music: trumpet (2 or 3), french horn (2-4), trombone (3), and tuba (1), more rarely cornet, natural horns and so-called Wagner tubas (→p.32). In the mixed make-up of classical chamber music written for strings and woodwinds, we usually find only the french horn from the brass family. However there are several types of ensembles consisting only of the brass instruments. In brass bands (usually military marching bands), we find trumpets, french horns and trombones, as well as cornets, alto horn, tenor horn, baritone or euphonium and tubas in their helikon or Sousaphone structural versions. Concert bands consist of a woodwind and brass mixture. In jazz and pop music, trumpets and trombones predominate, and in the traditional jazz styles such as dixieland, one also finds cornets, helikons and Sousaphones. The trumpet used in jazz is particularly narrow and short. Among the historical brass instruments, there is the valve-less trumpet and horn (natural trumpet and horn), the trombone and the cornetto (German: Hölzernen Zinken). Most instruments are produced in different forms and sizes.

Trumpets, french horns and other horns are so-called transposing instruments which means that the played tones deviate from the written notes. The following are non-transposing: the trumpet in C, the french horn in C, and the trombones and tubas.

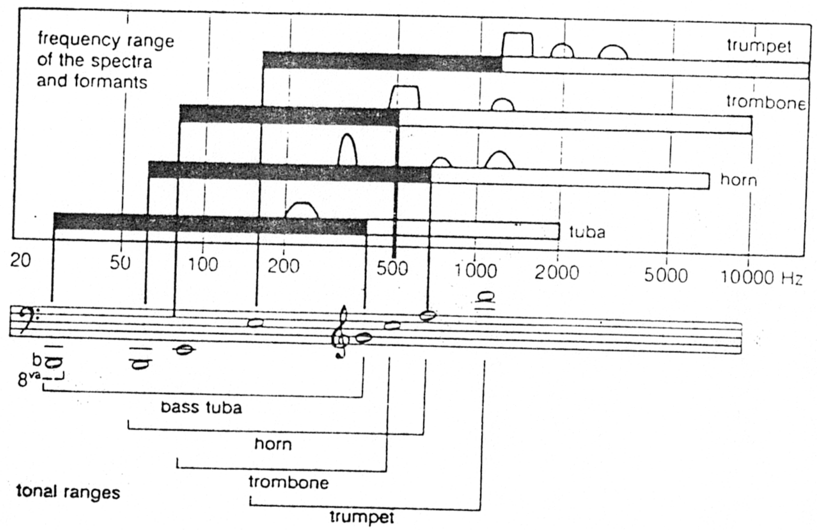
Sound acoustics

Sound generation: The brass instruments' mouthpiece serves as a support for the lips which, similar to the double reed, periodically interrupt the air stream. The interrupting frequency primarily depends on the resonant frequency of the air column within the instrument. The pitch is not influenced by finger holes at the side, as in the woodwinds, but by the tension of the lips, the air pressure and the extension of the instrument's length due to interchangeable, insertable tubing by means of valves or, as in the trombone, by means of a telescoping, u-shaped slide. The result is that, by contrast to the woodwinds, the entire sound emerges from the bell. This fact makes it possible to give an acoustical function to the brass bell which it cannot have in woodwinds, because there it only gives off the higher frequency components of each tone. In view of the fact that the bell of the brass instrument radiates all of the sound components, it is possible to optimize their sound energy transmission to the room. This increases the sound intensity of these instruments, which made them so useful as signal and fanfare instruments out-of-doors. The bell acts as an acoustical transformer and matches the low acoustical source impedance of the instrument to the higher acoustical terminating impedance of the room, thereby significantly raising the efficiency. A further improvement is provided by the great directional effect of the brass bell.

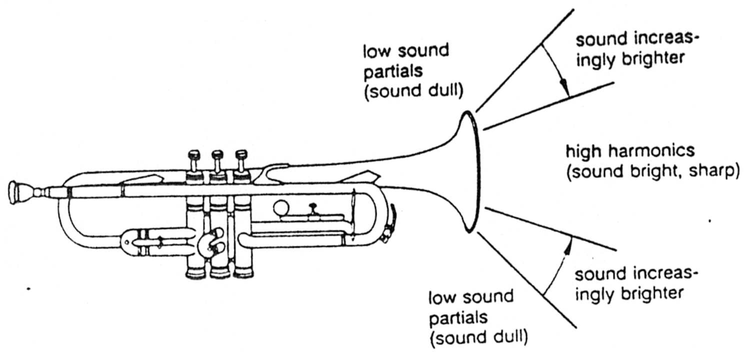
Attack: For a soft embouchure, the attack takes between 40 and 120 ms; for the trumpet up to 180 ms; for a sharp embouchure 20 to 40 ms; and for the french horn up to 80 ms. Typical for the attack of brass instruments is the so-called chiff which lasts on the order of 20 ms which contains predominantly harmonic contents below 1000 Hz. Too strong a chiff with an otherwise soft attack produces the well known squeal which is difficult to avoid, especially in french horns and historical brass instruments.

Instrument		Notation	Pitch referred to notation
Trumpets	trumpet in Bb (in C)	G-clef	a major second lower (as noted)
	(Instruments played today; player's choice.)		
	trumpet in G (in F, in E in Eb in D, etc.)	G-clef	a fifth (fourth, major third, minor third, major second) higher
	(Traditional instruments, but still notated that way today.)		
	bass trumpet in Eb (in C, in Bb)	G-clef	a major sixth octave, major ninth) lower
Horns	horn in F	G-clef, bass-clef (low parts)	a fifth lower
	(The instrument played today, also combined with the Bb alto horn.)		
	horn in C alto (in B, in Bb, in A etc.)	G-clef	as noted (a minor second, major second, minor third etc. lower)
	horn in C	G-clef	one octave lower
	(Traditional instruments, but still notated that way today.)		
Trombones	tenor trombone, tenor-bass tromb	tenor-clef, bass-clef	as written
	bass trombone	bass-clef, tenor-clef	as written
Tuba	bass tuba	bass-clef	as written
Cornet	cornet in Bb (in C)	G-clef	major second lower (as written)
Small horns	flügelhorn	G-clef	a major second lower
	alto horn in F (in Eb)	G-clef	a fifth (major sixth) lower
	tenor horn	G-clef	a major ninth lower
	baritone (Euphonium)	bass-clef, tenor-clef	as written
		G-clef, (brass band music)	a major ninth lower

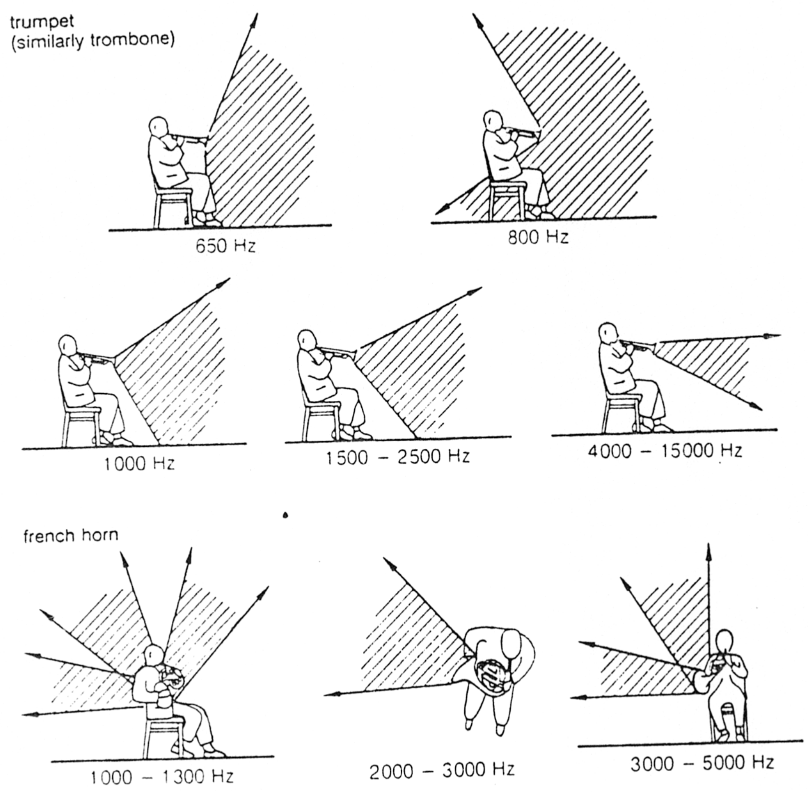
A. Notation of the brass wind instruments



B. Spectrum response ranges with formant positions and tonal ranges



C. Main radiating areas



D. Radiating characteristics

Pitch range (fig.B): The upper boundary of the pitch range is largely dependent on the skill of the player, just as in the string instruments. In the orchestra, the individual players specialize by assignment to a particular voice (e.g., 1st, 2nd, 3rd and 4th french horn); the higher voice to the 1st or 3rd, the lower to the 2nd or 4th horn.

Frequency range (fig.B): The widest tonal range occurs when playing fortissimo and depends on the pitch range of the particular instrument. The trumpet, as the highest pitch instrument, is capable of frequency components to 15,000 Hz; the french horn to 10,000 Hz; the trombone to 7000 Hz; while the lowest pitch instrument, the tuba, reaches but to 2500 Hz.

Formants (fig.B): The level increases in the formant regions are not as clear and in their frequency range not as even as is the case with the double reed instruments. Therefore their influence on the sound characteristic is lower.

Noise components: Noise content in brass instruments hardly play any role at all due to their very low levels.

Dynamic range and level: The dynamic range displays a pitch dependence (→p.40). For the trumpet the dynamic range gradually reduces from about 30 dB at the low frequency end to about 10 dB at the highest pitch tones. The french horn has a wide dynamic range of about 40 dB in its mid-range, reducing to about 20 dB at the upper pitch range. The trombone has the highest dynamic range among the brass instruments with values to 45 dB for certain tones. Brass instruments, therefore, have a rather uneven dynamic behavior. The level of low or high pitch tones, as with the dynamic range, is rather pitch dependent. For softly played instruments it increases by about 30 dB with increasing pitch. This means that low notes can be played very softly while high pitched ones can be played only loudly. The brass instruments are the loudest of the orchestra if one ignores certain percussion instruments. They are on average 5 to 10 dB louder than string instruments. Therefore it is not surprising that their number in the orchestra is smaller, there being some 60 strings but only 10 brass players in a large orchestra. The brass nevertheless can outplay the strings while the trombone achieves the highest levels.

Radiating characteristic

By contrast to the woodwinds, the directional characteristic of the brass is largely rotationally symmetrical about the bell and is therefore easier to manage. The radiating pattern becomes narrower with increasing pitch. The result is that the sound of a brass instrument becomes duller the farther off axis one moves (fig.C). This narrowing of the radiating pattern is not entirely regular. For the trombone, it widens once more at about 6000 Hz, for the trumpet at 800 Hz. While the principal radiating pattern for the trumpet and trombone points forward, the french horn directs it towards the back as a result of the way the instrument is held. The tuba directs its sound upwards. The directional pattern of the french horn is more complex than for the other instruments. It is divided into several angles (fig.D). The very strong directionality of the brass instruments, especially of the trumpet and trombone, allows the sound level to attenuate much more slowly in the direction of maximum radiation than in undirected other directions. This increases the reverberation radius for these instruments significantly (→p.26). The result is that, at greater distance from the bell, these instruments assume less of the room characteristics than do other instruments. Furthermore, the increased reverberation radius displaces the orchestral balances in favor of the brass with increasing distance, both as to instrument sound level and presence.

Instruments and their application

The number of percussionists in a symphony orchestra is relatively small when compared to strings and winds. For compositions from before 1800 there usually are only two timpani, in the 19th century often four or more with added bass and snare drums, cymbals, triangle, tom-tom, etc. In the 20th century the number, variety, and meaning of percussion instruments has been significantly enlarged. Added have been percussion instruments stemming especially from folk music and African, South and Central American and Asian countries. Some entered through the jazz world, others came directly into the orchestra.

In pop and jazz, the percussion group, also called the battery, has a very basic function. By contrast to classical music, it is constantly in action (beat). The standard outfit comprises at least one each of a bass drum, snare drum, at least two cymbals of different sizes, a foot operated cymbal pair (hi-hat), as well as two small tom-toms and one or more large tom-toms; add to that instruments depending on the style, such as bongos, congas, maracas and guiro, cow bells, wood blocks, gongs, etc. The piano is today's most universal instrument, appearing in all types of music both as soloist and in ensemble.

Sound acoustics

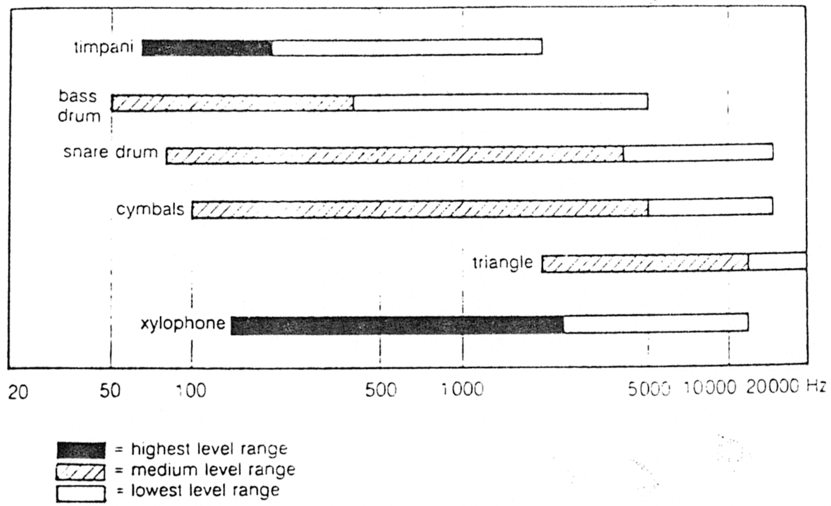
The sound of percussion instruments when compared to other instruments is more noise-like. This is especially true for the drums. Or it is marked by a certain out-of-focus tonality, caused by non-harmonic spectra as, for instance, for the triangle, gong and bells. The attack is extremely short because it is struck and this gives the battery its distinctive sound. Aside from the timpani and bass drum, the battery's spectrum extends to very high sound partials, often beyond the audible or reproducible range of 15,000 Hz (fig.A and B).

The timpani have a precisely perceivable pitch as the only such percussion instruments and are, therefore, written as music in the score. They are usually set to the required pitch by means of a pedal after first having had their membrane tuned. Timpani of different diameters are assigned different pitch ranges. The response only reaches to about 2000 Hz and the timpani, therefore, lack upper harmonics. The type of stick used (felt, foam, wood) determines the spectrum. Both the dynamic range and level of the timpani are extremely high.

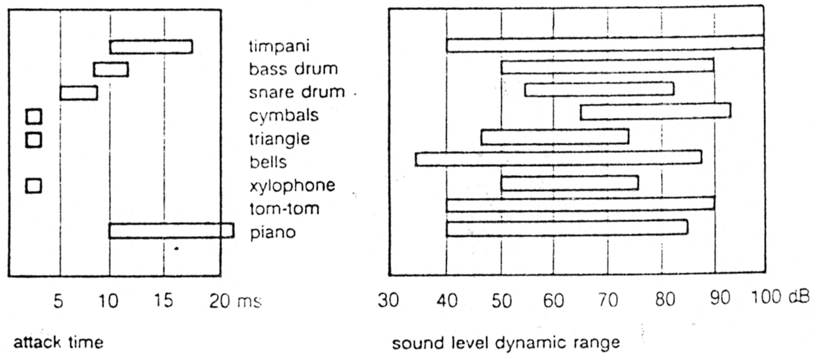
The big drum, in pop and jazz called the bass drum, has a very pronounced low frequency spectrum reaching, as it does, only to about 5000-6000 Hz. Due to the relatively low sensitivity of the human ear to low frequencies, the bass drum isn't perceived as very loud in spite of its tremendous level which exceeds that of the entire orchestra. For a close microphone position, the sound pressure level (SPL) is in the neighborhood of 100 dB. For pop music and jazz, the resonance membrane (head) opposite the playing one is usually removed and the tailing sound of the head damped through the use of carpeting and similar devices. The kick drum is smaller than the bass drum and therefore sounds quite different.

The snare drum, also called the side drum, has a comparatively narrow dynamic range and only a moderately high level. It may be played with snare springs against the resonant, opposite head. The frequency response range is again highly dependent on the playing method and may increase with loudness to as much as 15,000 Hz.

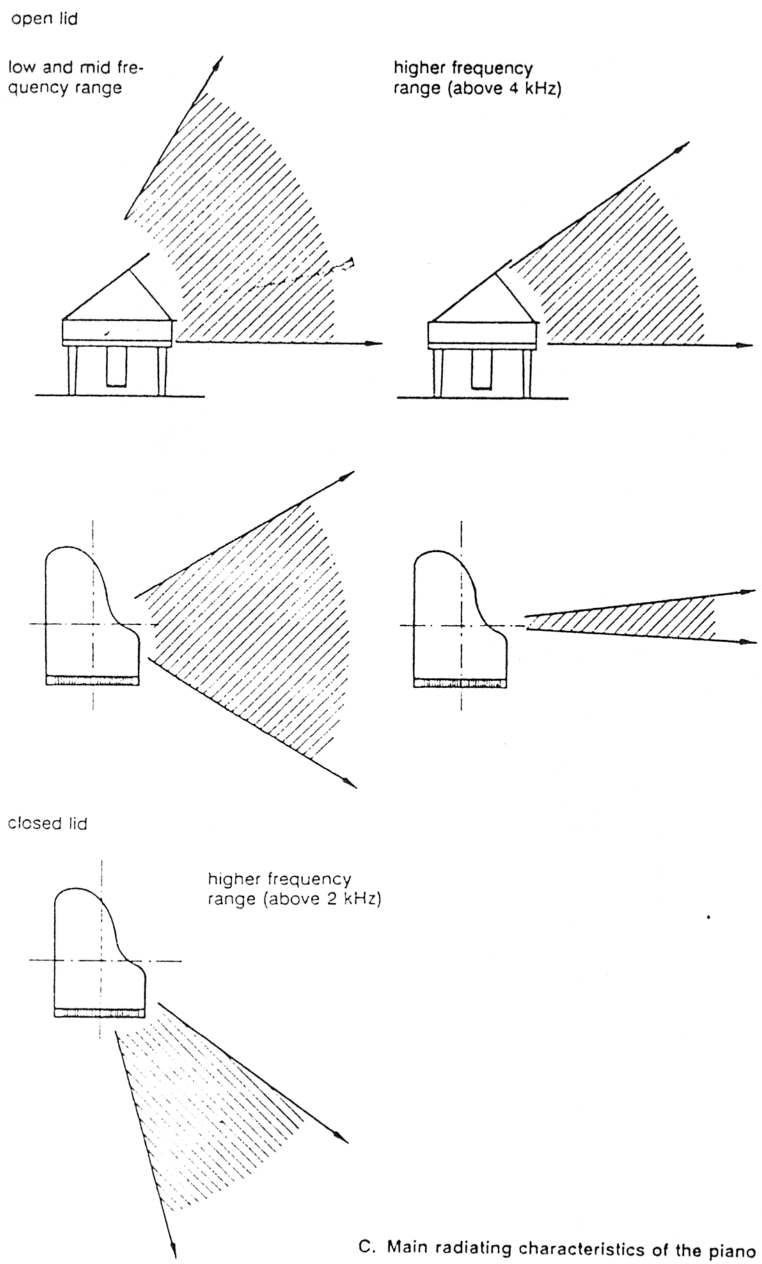
The cymbals have very sharp resonances up to 5000 Hz, size dependent. But even beyond that the frequency partials are quite strong. For brushed cymbals, the



A. Frequency response ranges along with the tonal ranges (timpani, xylophone) and areas of strong levels



B. Attack duration and dynamic range



C. Main radiating characteristics of the piano

resonances shift to just below 15,000 Hz.

The tom-toms have a tuned pitch which is not quite as recognizable as with the timpani. There are larger, free-standing tom-toms as well as smaller ones usually mounted atop the bass or kick drum. The congas are tuned similarly to the tom-toms, but they are not played with a stick but with the bare hand. The same is true of the smaller bongos that are likewise tunable.

The xylophone characteristically has an extremely short and noisy attack pattern. By contrast to the piano, the decay is on the order of the room reverberation at the low frequency end and in the high register even shorter. This probably explains the fact that, especially with the xylophone, the room tone with its information about the room itself, is especially audible.

Bells belong in the same family as the xylophone, e.g., the idiophones. These are mostly percussion instruments of vibrating plates or shells (bells, gong, tom-tom, cymbals, etc.), rods (xylophone, vibraphone, marimba, celesta, triangle, etc), tubes (bell chimes), or instruments with other shapes (wood blocks, castanets, etc). These instruments generally exhibit non-harmonic sound spectra. The bell exhibits a very special phenomenon known as the *strike sound,* i.e., at the moment the bell is struck, one hears a tone other than the one during the decay phase.

On the piano the strings are likewise struck and the attack is therefore rather short: 10-25 ms. The spectrum encompasses a range from 3 kHz to beyond 10 kHz. Formants are not sharply defined and mostly give a clue to the instrument's maker. Descriptive of the piano sound are the noise partials which are caused by the striking of the string but which die away quickly after that. These noises are centered between 200 and 1000 Hz. In the lower register they are largely masked by the harmonics of the tone but appear more prominently in the higher registers. The playing touch hardly has any effect on the attack but more on the timbre of the decaying tone. A special feature of the piano's spectrum is the spreading of its harmonic partials which may not be found in either string or wind instruments. This effect becomes clearer, the thicker and stiffer the string becomes in relationship to its length. Therefore, it is most noticeable for higher tones and for smaller size instruments, thus demonstrating the negative aspect of such small size instruments. In the most unfavorable cases the pitch sensation actually splits into a dual pitch perception. Since the piano tone is composed solely of attack and decay, it is the length of the decay which becomes most important. It is comparable to the reverberation time of a room (→p.36). The decay time, defined analogously to reverberation time, decreases with increasing pitch but, even within a tone itself, with increasing frequency of harmonics. Pianos displaying a bright, transparent sound, show a relatively short decay time of 20 s in the lower register as compared to the usual 30-40 s.

The radiation direction of a concert grand shows relatively wide angles both towards the side and upwards (fig.C). The radiating angle for high frequency components is surprisingly narrow. For a closed top, the higher sound components, which give the sound its presence are diverted towards the keyboard. As a whole, closing of the piano's top results in a duller sound with less presence but doesn't make it noticeably softer.

Speaking voice

Dynamic range and level: Human speech is rather soft when compared to the sound levels of musical instruments. As a guide, the following averaged maximum levels are valid for a microphone distance of 60 cm (24"); at twice this distance they are about 4 dB lower, and at half the distance 4 dB higher:

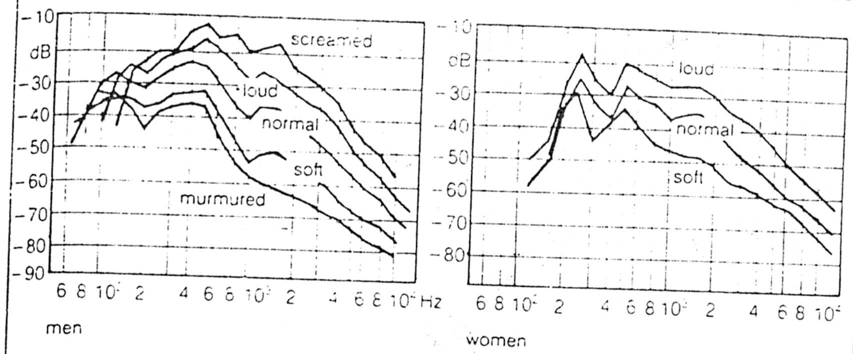
Averaged maximum levels	Normal speech softer	Normal speech louder	Loud speech	Average dynamics
Men	60 dB	65 dB	76 dB	16 dB
Women	58 dB	63 dB	68 dB	10 dB

If the speech sounds half as loud, this means a level decrease of 6-7 dB. Murmured speech lies another 5 dB below the level of soft speaking; very loud speaking about 5 dB above loud speech. The dynamic range for extreme forms of speech lies at about 25 dB for men and 20 dB for women. These values are valid for a so-called microphone voice, i.e., for a style of speaking which does not accentuate individual words, does not drop ends of sentences, and may be described as appropriate for an acting voice. In recordings, this sort of voice leads to a lower average level and with it, in practice, to a greater problem concerning the loudness ratio of speech to music in level metering.

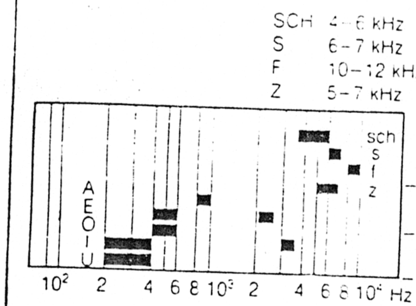
Level structure: The level structure is definitely impulse-like, e.g., strong, high level spikes caused by explosive sounds determine the highest level values. Short pauses between sentences, phrases, words, syllables, and phonemes interrupt the level pattern. This results in an average level which is far below the peak level, i.e., on average by as much as 12 dB or, in other words, at 25% modulation. The average level of popular music is generally assumed to be 6 dB below its peak content (at 50%) for serious music these values are 18 dB below or at 12%. Since the loudness approximates the average level, speech is perceived to be significantly softer than music when its peaks are at the same level. Therefore, it is impossible to achieve a speech/music loudness balance from a comparison of the peak level meter readings. Rather, the average levels derived from the peak levels must be used to achieve such balance. An approximate equal loudness between serious music and announcements is then likely if the music indicates peak and the announcements a level 6 dB lower. Naturally such numbers can only serve as an approximate guide; manner of speech and style of music play a very important part in this.

Frequency range: Fig.A shows the average spectra for male and female speech. The individual spectra are rather similar; however, with decreasing level, more and more high frequency partials fall below the threshold of hearing. The radiation of frequencies below 100 Hz (men) and 200 Hz (women) are largely independent of speech loudness; they mostly are dependent on the distance. Playback levels which greatly deviate from the loudness level at the microphone during recording produce disagreeable changes in the low frequency components of the voice; at high levels as boomy. It is the sibilants, particularly the S sounds with a range beyond 15,000 Hz, which have the greatest response range.

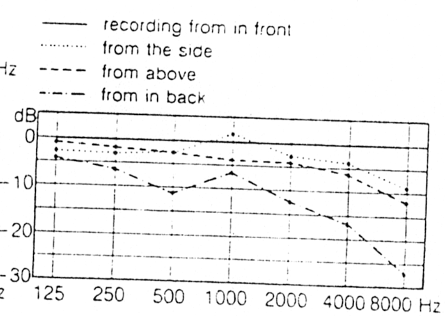
Tonal structures: The vowels are the musical components of speech; they have a melodic line spectrum. Various resonant-like peaks in the spectrum, the so-called speech formants, differentiate the individual vowels (fig.B). A response range of



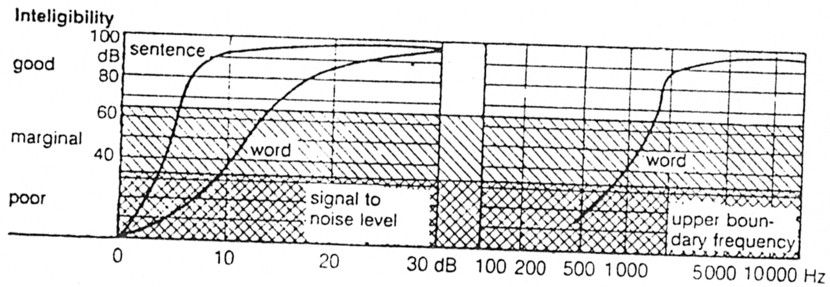
A. Average speaking voice spectra



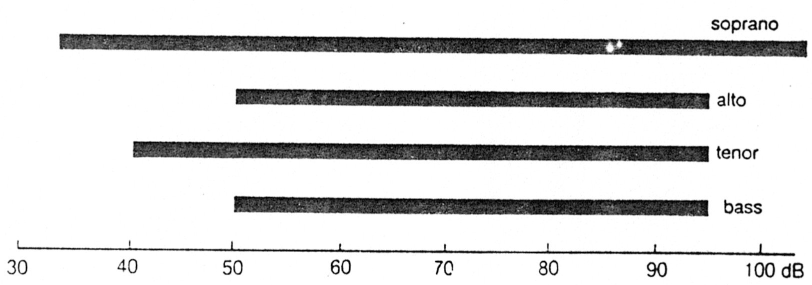
B. Speech sound formants



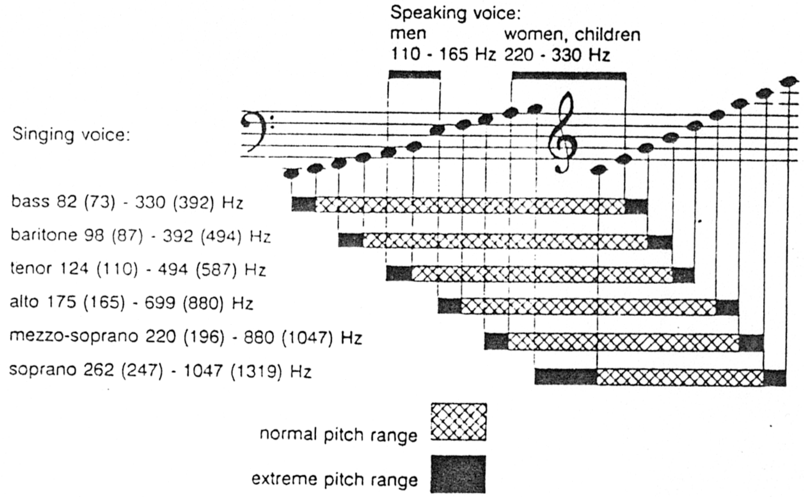
C. Sound coloration for various radiation directions



D. Word and sentence intelligibility with background noise and restricted band width



E. Speaking voice dynamic range and sound levels



F. Pitch ranges of the human voice (voice placement)

3600 Hz is required for their true transmission (the response of long distance phone lines). Vocal sounds such as L M N NG R are comprised of a line and a continuous noise spectrum. Explosive sounds (B P G K D T Qu) and fricative sounds (S SH CH F X Z) are pure noise spectra partially characterized by formant structures. The relatively high levels of S and SH above 10 kHz may lead to overload in sound storage and transmission systems due to the pre-emphasis used in them. The low level of the F sound may, by comparison, be easily overlooked during editing.

Radiating characteristic: Due to the directional character of the human voice, its level is about 5 dB lower towards the side, and 10 dB lower at the back of the head. For a microphone level identical to that used in frontal recording, this means an increase in diffuse sound level of 5 dB, respectively 10 dB when recording towards the side or back of the head. The character of the sound is modified significantly, i.e., frequencies above 1000 Hz are radiated to the side and the back at lower level (fig.C). The greater reverberation portion and the decrease in higher frequency sound components have an effect similar to a considerably greater microphone distance in a frontal recording. In rooms with little diffuse energy, and especially for close placement of the microphone, a position of the microphone to the side may help to avoid the overload created by sibilants.

Intelligibility: Fig.D shows the word and sentence intelligibility as a function of the weighted signal-to-noise ratio and the upper limit of the transmission range. Practically 100% word intelligibility is assured for an upper frequency limit of 5 kHz and a weighted signal-to-noise ratio of 30 dB.

Singing voice

The acoustic differences between the singing and speaking voices are not as great as would appear from the aural perception. Since only sounds with an harmonic structure possess pitch, these are particularly accentuated and expanded when singing. While the speaking voice pitch changes gradually and often, the singing voice is tied to certain pitch steps. The formant positions are matched to a certain extent to the pitch of the fundamental that generally lends the vocal character a certain darkening effect. The co-called singing formant between 2800 and 3000 Hz is of considerable meaning for the timbre of the voice. This is generally connected with an overall amplification of the higher frequency sound components which then give the human voice its ability to be heard even over a loud orchestral background. This does not take place in speech. A hallmark of the trained voice also is its vibrato and tremolo. An especially pronounced accentuation and stretching of the vowels, singing formant, vibrato and greater loudness and dynamic range define, purely acoustically, a clearly well trained voice.

Dynamic range and level: Both dynamic range and maximum level of the singing voice are surely a function of the particular singer and type of music. High female (soprano) and male (tenor) voices attain the widest dynamic range. The peak level of a soprano for a normal microphone distance may reach over 100 dB SPL (fig.E).

Vocal ranges and character: The pitch range of a human singing voice determines its vocal range (fig.F). Soprano, alto, tenor and bass are the main ranges. The suitability of a voice or soloist for a specific role is described by its character (e.g., dramatic soprano, coloratura soprano, lyric alto, Heldentenor, young lover, basso buffo).

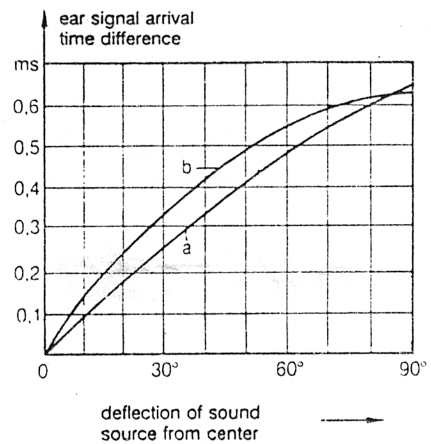
Spacial hearing

The ability to hear spatially results from the interaction of directional and distance perception. For directional perception in the horizontal plane it is the difference between the signals arriving at both ears that are determining, while directional perception in the vertical plane depends on the sound coloration which is a function of its vertical angle. For the judging of distance, it is the loudness level, timbre and the ratio between direct and diffuse sound which are determining. The experience and practice of the hearing mechanism has a significant influence on the localization accuracy. For presentation of stereophonic sound images via loudspeakers or an earphone pair, our hearing mechanism uses acoustic properties different from those used in natural hearing.

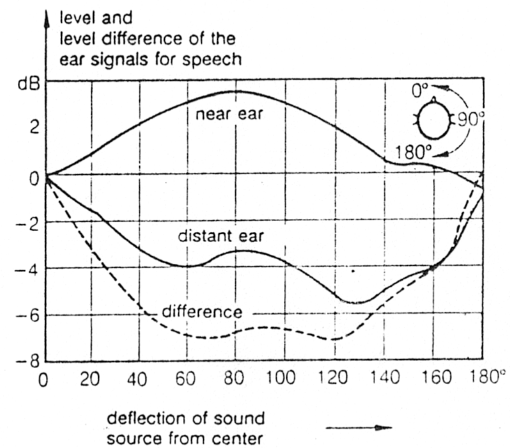
Spacial perception of a sound source

If a sound source is located in the horizontal auditory plane off center, there appear time-of-arrival differences up to a maximum of 0.6 ms (fig.A) as well as frequency dependent level differences and, with these, timbre differences between the two signals arriving at the two ears. For the speech spectrum the level differences amount to about 7 dB (fig.B). The same amount applies to music as well, depending on the particular spectrum. In general both the time-of-arrival and level difference act in concert to impart localization. For sine wave signals, the level difference is a function of the frequency: below 500 Hz there are virtually no level differences, while they increase with rising frequency above 500 Hz. The time-of-arrival difference of sine wave signals is sensed as a phase difference between the ears. They impart a unique location only up to 800 Hz; above that multiple locations. All in all, the localization of sine wave signals is difficult and often leads to disorientation. The localizing accuracy in the horizontal plane depends on the type of signal. Broad band and impulse type signals are localized most accurately (about ± 1°), narrow band signals least accurately (about ± 5° to 10°). Narrow band signals may lead to multiple localization if the signal is of short duration or if the head makes no direction finding movement. Steady tones are mostly localized at low frequencies through time-of-arrival differences and for higher frequencies predominantly through level differences. For broad band steady signals, it is primarily level differences which determine the sound source direction but also the time displacement of the signal envelope at the ears.

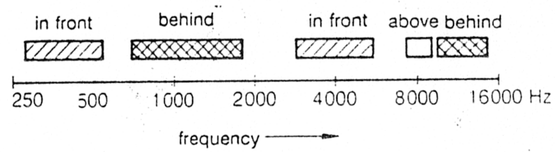
If a signal moves upward in the vertical auditory plane, the level and time-of-arrival differences remain basically unchanged. The information about the rising sound source is mainly taken from the change in timbre caused by the diffracting effect at the head and ear lobes (pinnae). Sounds emanating from different directions obtain direction specific boosts in certain frequency bands, i.e., the so-called direction determining bands (fig.C). Narrow band signals cannot have such complex spectral changes and therefore localization of such signals becomes difficult if not impossible. The localizing accuracy in the vertical plane is far inferior to that in the horizontal plane; it depends clearly on the familiarity with the sound source and on the experience of the hearing mechanism. Unknown speaking voices may be localized to ± 15° to 20°, known ones to ± 10°. White noise may be localized as accurately as ± 4°. Localization in the vertical plane is not possible for narrow band signals; the signal direction is determined from the signal frequency.



A. Time-of-arrival differences between the signals at the two ears
a. calculated
b. by experimentation with headphones

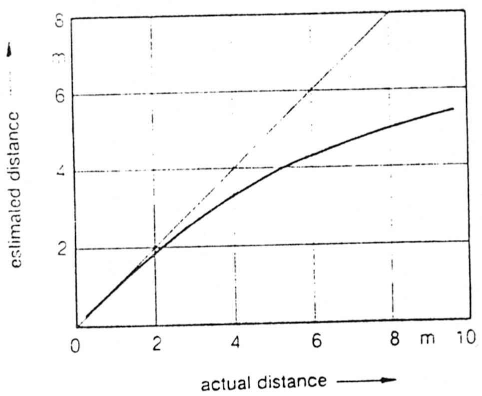


B. Level differences between the signals at the two ears for speech

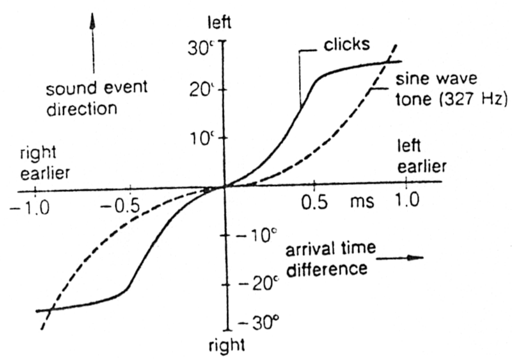


C. Direction determining bands for sensation of the vertical plane

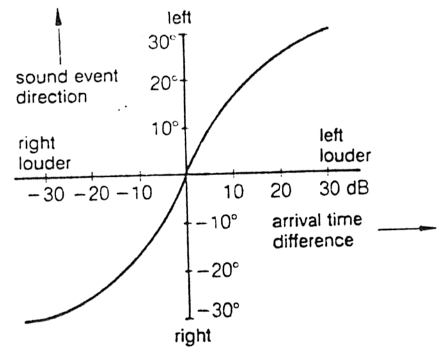
D. Connection between the actual and estimated sound source distance (announcer in minimally reflecting room)



E. Delay difference between two loudspeakers for stereo reproduction and the resulting deflection of the phantom sound source



F. Level difference between two loudspeakers for stereo reproduction and the resulting deflection of the phantom sound source



Distant hearing

The ability to hear distant sounds is based even more on auditory experience than is the hearing in the vertical plane. For distances between 3 and 15 m (10 and 49 ft) out-of-doors, it is the decrease by 4 to 6 dB for every doubling of the distance which provides the best clue to the sound source's distance. For greater distances it is the attenuation of the higher frequencies through air absorption which plays the most important role and in the near field the linear distortion of the spectrum resulting from diffraction around the head. Distance sensation is improved markedly in enclosed spaces through evaluation of the ratio of direct to diffuse sound, permitting recognition of even small distance differences. Altogether the recognition of the sound source's distance underlies very complex factors making it difficult to give generalized rules. On the one hand, the perceived accuracy is quite poor, and, on the other, increasing distance of the sound source results in underestimation of the distance (fig.D).

Spacial hearing with loudspeaker reproduction

In stereophonic reproduction the spacial distribution of auditory stimuli is enabled through a special phenomenon of sound perception which is of no significance for the hearing of natural sound sources – the creation of so-called phantom sound sources. A phantom sound source is created exactly in the center between two stereo loudspeakers when both speakers give off identical signals simultaneously. For increasing time delay and/or level difference between the loudspeakers, the phantom sound source moves away from the center until it appears to come to rest in one of the loudspeakers. The louder or earlier arriving signal determines toward which loudspeaker the phantom sound source moves. The required delay time for a particular angular deviation (fig.E) depends somewhat on the type of signal. A time-of-arrival difference of only 1 ms results in the greatest possible deviation of 30° from the center for the usual stereo loudspeaker reproduction setup in an equilateral triangle with the listener. Level differences (fig.F) lead to more stable phantom sound sources. In practice, level differences of 15 to 20 dB permit phantom sound sources to deviate so far to the side that they become stable, real sound sources located at one of the loudspeakers. It is a condition for the formation of phantom sound sources that the listener be located on the exact center line between the speakers. Even for small deviations in location the phantom sound source wanders towards the nearer loudspeaker. The so-called stereo horizon within which true localization of the phantom sound source is possible is only about 20 to 40 cm (8" - 16") wide for normal stereo setups (→p.89).

The theory of summation localization attempts to explain the spacial perception of stereophonically reproduced sound sources. It assumes that for the perception of phantom sound sources in loudspeaker reproduction, summed signals result from the superimposition of the sound fields of both loudspeakers, that are not further differentiated by the hearing mechanism. They are not identical, but are equivalent to those signals of a natural sound source located at the phantom sound position. By contrast, the association model goes back to prior experiences of the brain. In two recognition processes which run sequentially, we first determine the sound source location and then the nature of the sound source. For instance, those sound colorations which provide directional information for a raised sound source or that result from the superimposition of time displaced stereo loudspeaker signals (comb filter curve) are canceled. By contrast to the theory of summation localization, the association model allows the unchallengeable explanation of spacial hearing.

The recording of string instruments

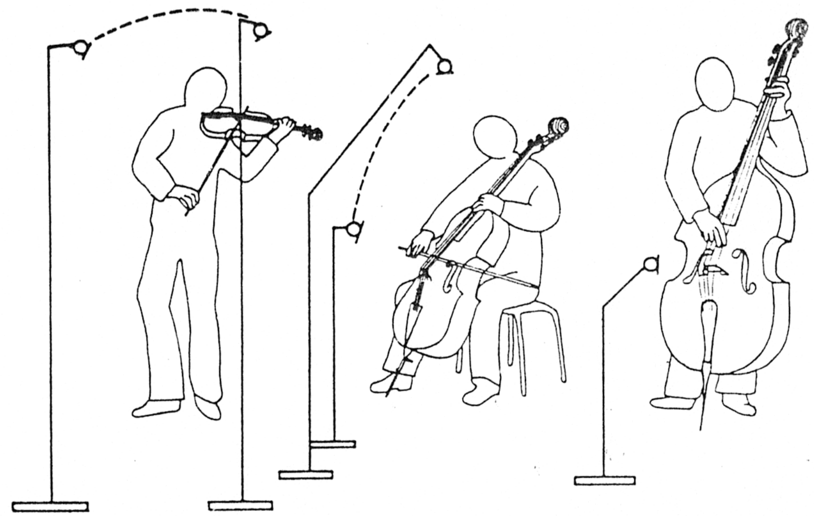
Directional characteristic, formants

The directional characteristic of string instruments (→p.44), their principal attribute for the setting up of microphones, is characterized by less definition but greater complexity compared to the wind instruments. The strings' timbre changes less with the microphone placement, therefore, the microphone setup is not as critical as it is with wind instruments. Recording in the near field results in a curve comparable to a comb filter at higher frequencies due to irregular radiation (→p.44). With the addition of artificial reverberation, this curve is also transferred to the frequency response of the reverberation. A greater sound source to microphone spacing helps to avoid this effect which leads to a sound shrillness. The main formant range (→p.44) is important for reproducing the timbre of the particular instrument. The formant range, however, is clearly defined only over a relatively small radiating angle. In the violin the sound determining A-formant (about 1000 Hz) is principally radiated by the top of the instrument's sound box, so that the most favorable microphone location is in this area as well. The microphone distance largely depends on the room acoustics, the recording technique used and the æsthetics of the recording (fig.A). The viola displays similar characteristics to those of the violin. The violoncello radiates its low frequencies principally in a direction at right angles to the front of the instrument. It is directly in front of the instrument that it has its fullest sound. For a higher positioned microphone, the sound becomes restricted since those frequencies which impart that kind of sound quality (2000 - 5000 Hz) are radiated upwards (fig.A). As a result, the celli often sound less full in the orchestra than do the violins. With the contra-bass as well, it is best to aim the microphone at the body's front, but not at the f-holes. Typical for the sound of the double bass are the bow hair noise oscillations which may reach frequencies up to a range of about 10,000 Hz. Their buzzing gives the sound of the double bass a characteristic presence which may be heard at all times over the sound of the orchestra. The otherwise poor localizability of the double bass is improved significantly by this buzzing.

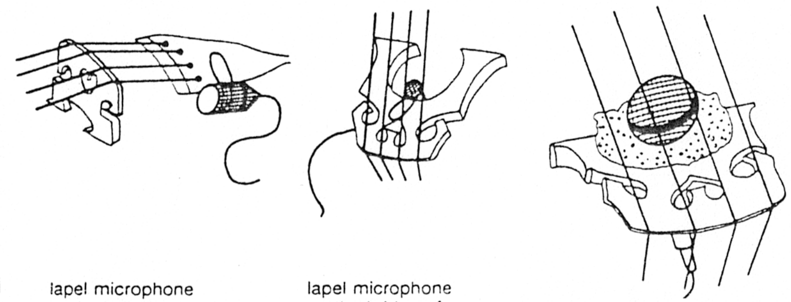
Extremely close recording

An extremely close placement of microphones to string instruments gives a very present and dense sound image even in the presence of a great deal of reverberation. This sound is not a natural one and is, therefore, suited largely to the recording of popular and similar types of music. In the presence of simultaneous sound reenforcement, the relatively low resulting microphone amplification improves the feedback characteristics. Such a method may also be desirable on TV sets for optical reasons. Lapel microphones are usually fastened to the tail piece or are aimed at the instrument body. Such microphones may actually be clamped to the bridge of celli and double basses due to these instruments' larger size (fig.B). Lapel microphones display no low frequency rise in the near field due to the fact that they are normally pressure transducers.

The fact that lapel microphones are often lavaliere microphones with their particular response curve is of no significance (→p.74). For extremely close microphone placement one may also use general studio microphones on stands but this demands cooperation from the musicians in view of the limits placed on their movements. For celli and double basses, the microphones, wrapped in foam, may be clamped under the bridge (fig.B). Vocalist and pressure microphones are preferred in this application. Beyond that one may want to consider boundary surface microphones without their mounting plate or even contact microphones.



A. Microphone placement appropriate for string instruments



B. The use of lapel and close-up microphones for the recording of string instruments

Directional characteristic, formants

When compared to the strings, wind instruments are very directionality dependent in their radiation (→p.48,p.52). For the recording this means that the choice of microphone position provides greater influence on the instruments' timbre than does the choice of microphone type. Often even a small change in the location of the microphone or musician will result in significant changes in sound, and there are microphone positions which will result in an unacceptable sound. In practice there are two significant differences in the recording of woodwind and brass instruments.

Brass instruments, by contrast to the woodwinds, have a large bell which results in a much greater directional concentration, even in the mid-range. On the other hand, this larger bell provides for a greater acoustical coupling of the instrument to the room resulting in greater sound energy radiation. Therefore, brass instruments may be significantly louder than woodwinds. They radiate about 5 to 10 times the sound energy meaning that they sound twice as loud as the woodwinds. The woodwinds, unlike the brass, have finger holes along the length of the instrument, which radiate significant portions of the sound. As a result, the directional characteristics of woodwinds are more complex and are not rotationally symmetrical about the end bell.

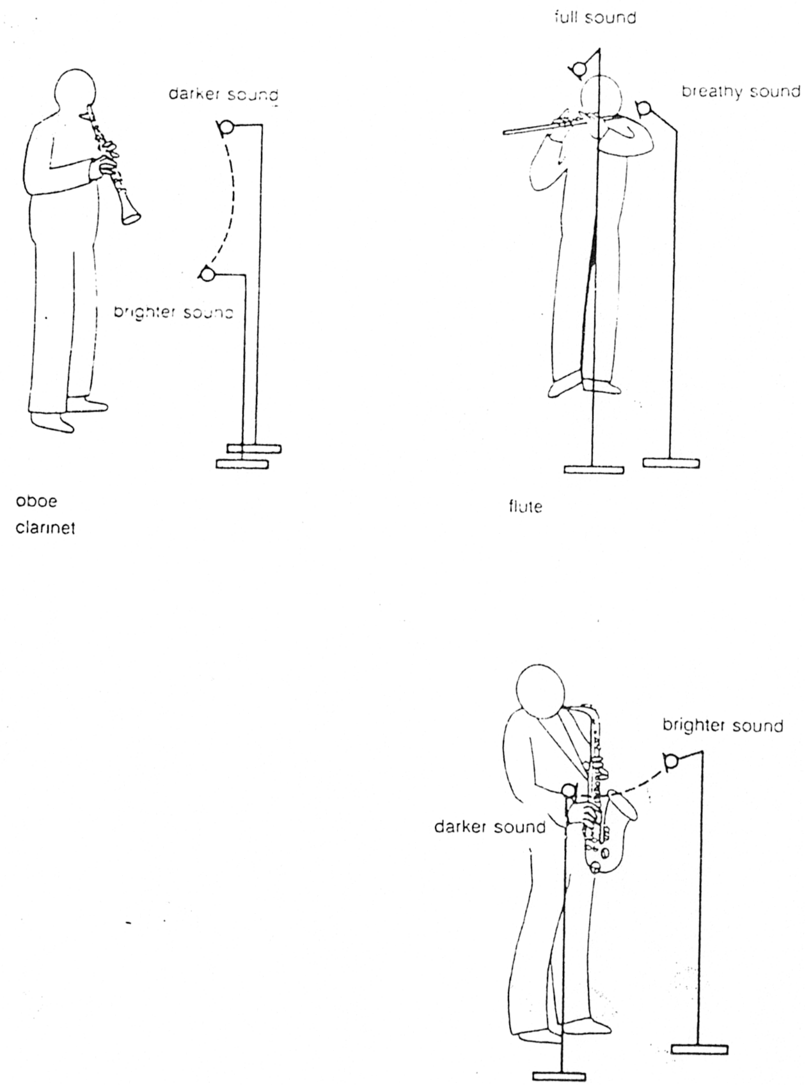
The bell direction of the instruments also is not uniform. They may point at the floor (clarinet, oboe, soprano saxophone), more or less horizontally forward (trumpet, trombone etc), horizontally towards the side (flute), upwards (bassoon, tuba) or behind the player (french horn). For this reason there result different microphone positions for the individual instruments as well.

It is rather strenuous to play a brass instrument, and, therefore, it is important that one spare these musicians' strength.

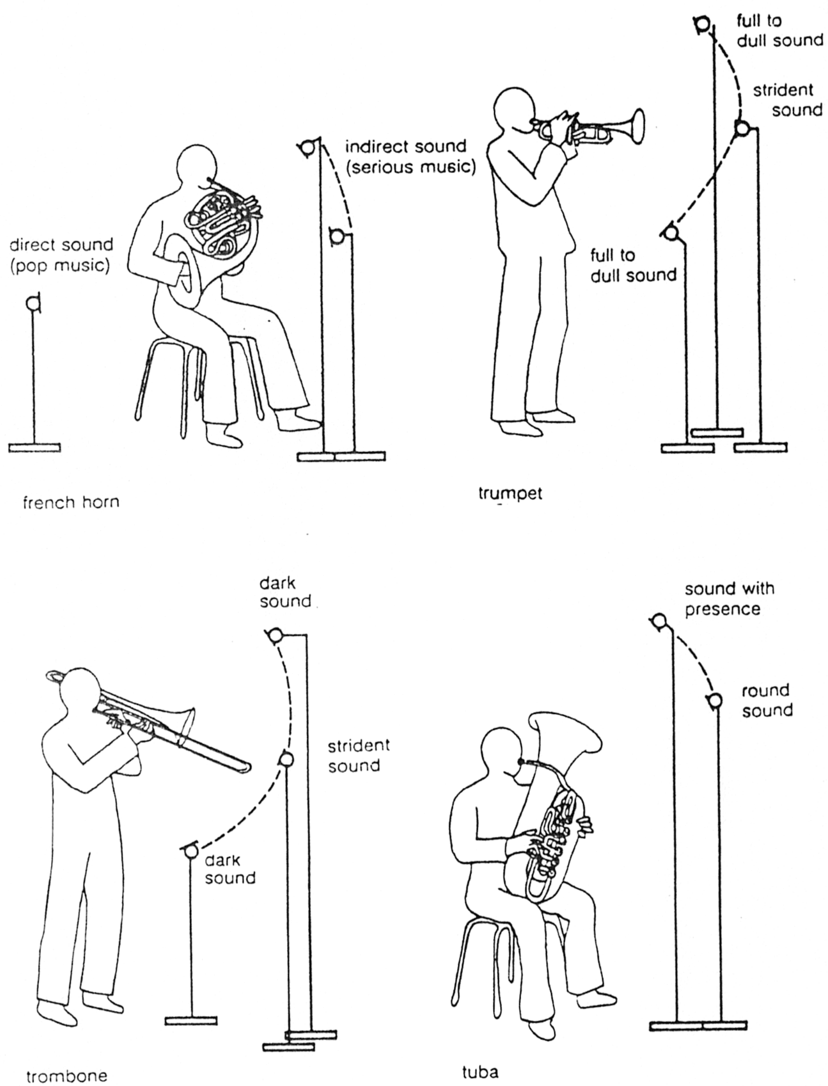
Woodwind instruments

For the clarinet and oboe, sound components up to about 3000 Hz are predominantly radiated through the finger holes. Higher components radiate from the bell (→p.48). This permits the selection of a brighter or darker sound through proper orientation of the microphone (fig.A). In the area directly in front of the bell, the sound is unnaturally tight and strident, and that is why this usually is not a microphone position to be recommended. The microphone spacing is determined by the recording technique employed. The closer the microphone, the greater the effect of even slight motion of the instrument on the sound timbre. For this reason distances of less than 50 cm (20") are recommended for use only with experienced studio musicians who are able to keep their instruments in the predetermined position. If the microphone is placed close to the mouthpiece, one obtains the disagreeable effect that each tone seems to come from a different distance, depending on which keys happen to be open. Normally the microphone is directed downward from above the instrument; in special cases, e.g., to achieve better acoustical separation from neighboring instruments, it may be directed at the instrument from the side, without any adverse effect on the sound quality. The occasionally heard noisiness of the keys appears only with instruments that are badly worn.

With the flute a microphone position over the keys yields satisfying results (fig.A). The sound in front of the instrument's end (it has no bell) is weak, noisy and tight. Such a position is not recommended in practice, however, for pop and jazz, where breathiness is often included in the musical intent, a recording made directly at the mouthpiece may be desirable.



A. Microphone placement for woodwind instruments



B. Microphone placement for brass instruments

The saxophone, whose bent-back end bell faces the keys themselves, permits the recording of both the sound partials radiated from the finger holes and the bell simultaneously. For microphone positions to the side, the sound becomes warmer and gentler. Positions extremely close to the end bell result in a dull and tubby sound. For the soprano saxophone with its in-line bell, the same parameters apply as for the clarinet.

The bassoon radiates high frequency sound partials diagonally forward and upward, the low frequency sound components, as in all woodwinds, to the side. The recommended microphone position is the same as that for the oboe and clarinet.

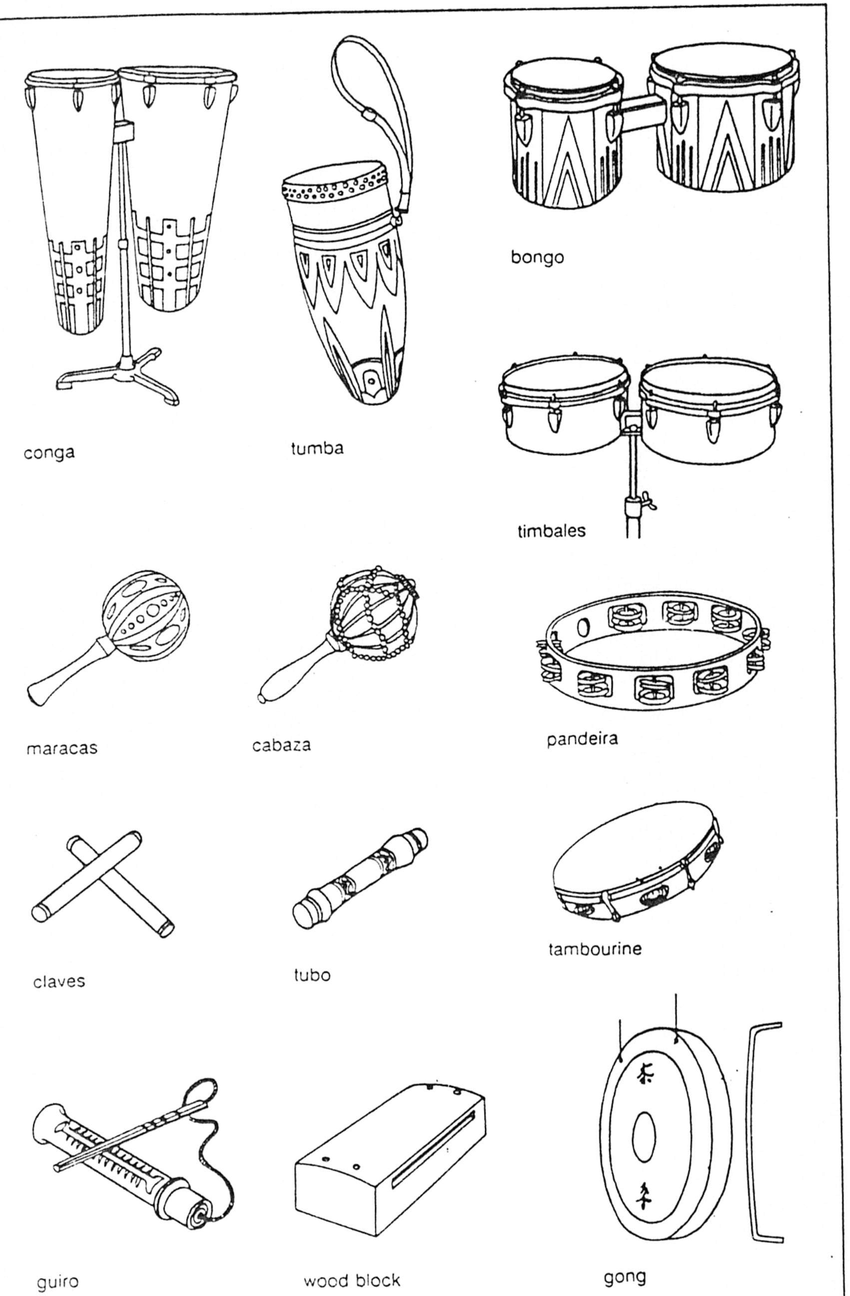
Brass wind instruments

The french horn is the only orchestral instrument which radiates its sound behind the player (→p.55). Therefore, the french horn sound in the orchestra is always indirect and rather spacious. Since this position of the instrument has remained unchanged since it entered the orchestra, composers have used the french horn in a way which corresponds to its unique sound. It is used as a sound integrating instrument often intended to sound as if coming from a great distance. That is why a microphone position in front of the bell is not recommended in serious music or for chamber music (fig.B). Sound reflecting surfaces behind the french horn lead to disturbing reflections. For the multi-microphone recording of popular music the french horn must also be recorded in the near field, e.g., the microphone must be placed in front of the bell. For semi-classical music, the setup used for serious music might be applicable. Even though the nonlinear distortion in magnetic recording has been reduced markedly, it is nevertheless the french horn which produces the most audible distortion at full level. Similarly sensitive are the recorder, trombone ensembles, and children's chorus.

The trumpet and trombone may be treated similarly when it comes to recording (fig.B). The tone is the brightest on axis but, by contrast to the woodwinds, sounds fairly agreeable. As one deviates from on-axis, they become increasingly dull (fig.B). Both instruments achieve the orchestra's highest loudness levels, with the exception of a few percussion instruments. Sound pressure levels in excess of 120 dB are obtained at a distance of 50 cm (20") in front of the bell; at 20 cm (8") over 135 dB. This places their sound pressure levels into a range that cannot be reproduced distortion-free by any but the most modern professional capacitor microphones. Significant here is the overload sound pressure given in the microphone's technical data for 0.5% total harmonic distortion.

The tuba, the bass instrument of the brass, radiates its sound upwards like the bassoon, the bass instrument of the woodwinds. The aggressive attack which typifies the tuba's sound is best obtained from a microphone position above the instrument. Moving more towards the side causes a more rounded sound while embouchure sounds decrease (fig.B). The Sousaphone is a tuba which radiates its sound forward and is, therefore, most often used in marching bands.

The other horns radiate their sound in various ways depending on their construction. If they have a trumpet shape (Flügelhorn, alto horn) forward; in their Waldhorn shape (alto horn) towards the back; and in their tuba form (alto, tenor, baritone horns) diagonally upward.



C. Less frequently used percussion instruments

In the bass or kick drum, the drum head opposite the one being played is generally removed, permitting the microphone to be placed about 10 to 15 cm (4"-6") behind the playing head. This provides a dry sound and good separation from the other microphones. A blanket or foam material, possibly weighted down with a rock, is often pressed against the drum head to obtain an even drier sound. The recommended microphone is a dynamic one which is not proximity effect compensated (→p.73). The bass drum has by far the highest level in the spectrum of a recording and, therefore, determines its maximum level. That means that the bass drum balance can only be increased at the expense of other instruments or voices. A desirable reduction of its level without influencing its sound quality may only be achieved by filtering out its lowest sound components below about 100 Hz. An electrical suppression of the sound hang-over is possible through use of a noise gate or expander and the furthering of a rapid, precise attack is obtainable with a low threshold set limiter. The side or snare drum is picked up at a distance of 5 to 10 cm (2"-4") from above (fig.B). Important here is the lowest possible cross talk from the cymbals. For the tom-tom the microphone is suspended over the drum head just as in the snare drum. One microphone is well able to handle two tom-toms. The resonance drum head may also be removed from the suspended tom-toms. Then the microphone may be inserted from the bottom into the inside of the tom-tom (fig.B). To effect an even greater damping of the snare drum or tom-tom skin, one often attaches a felt or foam strip to the edge of the skin with adhesive tape. The cymbals microphones are suspended 30-50 cm (12-20") above the cymbals while for the hi-hat a lesser distance is permissible.

Additional instruments to those listed, which form the nucleus of the percussion group, may be added (fig.C). Gongs are recorded from behind at a distance of 30-80 cm (12"-32"). For closely spaced microphones, each gong gets its own microphone. Bongos and congas as well as the tumbas (similar to the conga) are usually found in pairs. If they are important then it is best to use two microphones per pair to allow a stereo width expansion of the sound source. The individual microphones of the percussion group are usually corrected through the use of filters or equalizers. For the bass drum it is the high frequencies, for the cymbals and hi-hat it is the low frequencies that are filtered out. Microphone setup and console settings generally take quite a bit of time and care.

Jazz and folk music

Jazz recordings may be made using the same multi-microphone technique used in popular music. If a well balanced sound reenforcement system is needed as well, then only this method is applicable. Often the techniques applied in serious music recording are used. Then one finds methods using main and support microphones, or individual microphones set up at a somewhat greater distance. The method to be used largely depends on the musical style and must be selected in consultation with the musicians. This is also true for other types of music which fall into the area between jazz and pop.

For recording folk music etc., the same things hold true that have been stated for jazz. A method is often employed which falls somewhere between multi-microphone and support microphone techniques.

Serious music

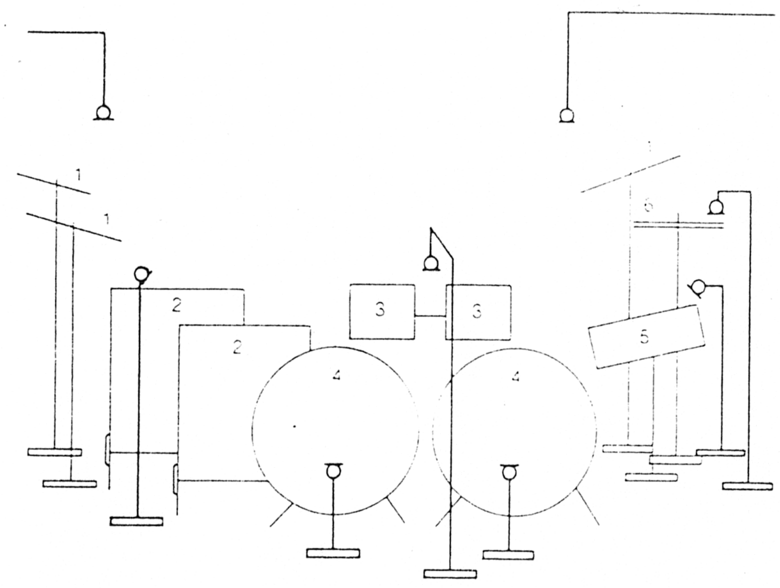
Far into the 19th century music accorded only a minor role to the percussion instruments (→p.56). Only two or three timpani are noted regularly. From their musical function, they act as the bass fundamentals of-the trumpets. It is only during the 19th century that this connection was dissolved. In order to record the timpani with tonal precision, it is recommended that a support microphone be set up in the near field, even for spatially oriented techniques such as XY, MS and AB. For judicious mixing this permits a high degree of sound presence with a slightly earlier attack than the orchestra's.

The triangle does not require its own microphone. It is supposed to provide a general, diffuse brightness to the orchestra, comparable to the Zimbelstern of the pipe organ. The various drums are only then given their own support microphone if they have some specific function, something one finds in many works of the 20th century. The same thing is true for the cymbals, which display very high frequency sound components, and which, for a fortissimo, may raise the entire level of the orchestra considerably. This also holds true for the tom-tom. The support microphones for these instruments are generally at a height of 1.5 m (59"). At this distance only cardioid or hyper-cardioid patterns are conceivable.

Semi-classical and popular music

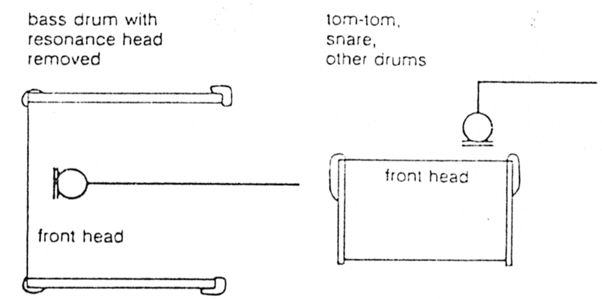
One finds the most varied possibilities in the recording of percussion (drums). The method to be used depends not only on the type of music, the playing technique, and the acoustical and technical situation, but also on the sound to be achieved and, therefore, on taste trends. The range of possibilities extends from extremely near field microphones for every instrument to a stereo microphone in intensity stereo or two individual microphones for the entire drum set.

The most usual contemporary method of recording drums is to use a single microphone for every one or two instruments (→p.94) and two overheads in addition; two microphones 50 cm (20") above the cymbals, or a stereo microphone above the entire arrangement. The stereo location assignment of the individual microphones or instruments corresponds approximately to the natural seating of the instruments. This is needed in any case when using overhead microphones in order to avoid double imaging on the stereo horizon. But aside from this, the setup of the percussion instruments (fig.A) provides an advantageous imaging for stereo: a) low-frequency instruments (bass or kick drum, tom-toms) in the center because they are difficult to localize anyway and because they would otherwise waste groove space on phonograph records; b) high pitched instruments (cymbals, hi-hat) at the sides of the stereo horizon because they make the stereo width especially clear. In general, the microphone distance is or may be a bit greater in pure studio recordings than on stage with simultaneous sound reenforcement. As to microphone type, dynamic microphones are often used for subjective reasons or because of their robust construction, their virtual freedom from overload and the fact that they require no powering. Because of the fact that modern day capacitor microphones are equally robust, that their overload reserve is usually sufficient, and even the simplest of consoles now offers phantom powering, it is only the subjective sound judgement which should determine the type of microphone. Due to the high sound pressure levels encountered, it is recommended that the overload protection switch be used.



A. Microphone placement for percussion instruments

1 cymbals
2 standing tom-toms
3 hanging tom-toms
4 bass drum
5 snare drum
6 high hat



B. Microphone placement for drums

Classical and acoustic guitars

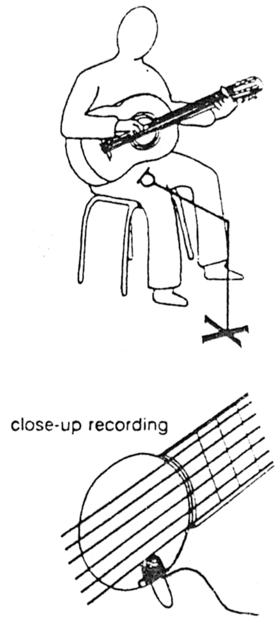
Since the classic guitar has a rather soft sound, the microphone spacing must be relatively close, even in serious music. About 50-100 cm (20"-40") is usual, otherwise the acoustical and electrical signal-to-noise level ratios are too small. The microphone is preferably to be aimed at the area below the bridge (fig.A). This generally applies also to the acoustic guitar, a generic expression for the various forms of the guitar without electrical amplification as used in the popular music field. There are problems caused by this instrument's low output, especially if a sound reenforcement system is in operation during the recording or if the guitarist sings and plays at the same time, requiring careful balancing. In such a case a lapel microphone mounted to the sound hole, may bring the desired results. In those guitars which also have a built-in pick-up, this is often mixed in with the microphone. In such a case it is the pick-up which predominantly supplies the mid frequencies while the microphone adds the low and high frequency components, so that it would be feasible to reduce the microphone's mid-frequencies by means of an equalizer to help the feedback problem in sound reenforcement.

The same technique that is valid for the guitar also holds true for the mandolin and lute. To pick up the harp, the microphone should be directed at its sound box. Harp pedal noises usually stem from worn felt padding around the pedal slots.

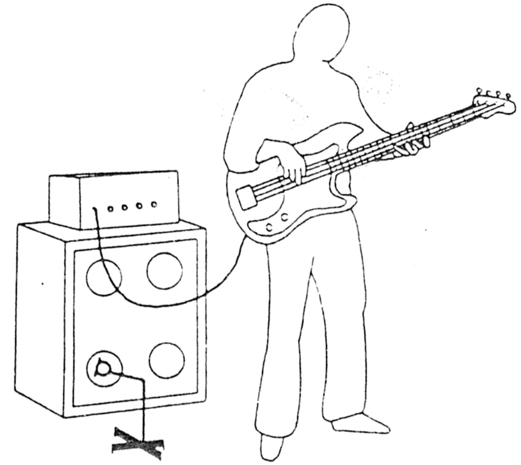
Electric guitar

Only the amplification and reproduction of the guitar via loudspeakers has permitted the guitar to become one of the most important instruments of modern popular music. For one thing the amplification offers an opportunity to influence the sound in many different ways using effects units, usually foot pedal controlled (fig.C). Naturally the pick-up and the guitar amplifier as well as the construction and materials of the instrument and strings have a significant influence on its sound as well. Dynamic pick-ups must remain well away from magnetic fields such as a power isolation transformer that might be located under the musician's chair. Since the guitar should sound the way the guitar player adjusts it, a pick-up with a microphone in front of the guitar amplifier is the preferred method (fig.B). The microphone is to be set up in the near field directly in front of the guitar amp loudspeaker facing the center of the cone. For non-coaxial, two-way systems, two microphones are needed. If a directional microphone is used at this close a distance, a bass roll-off must be provided either by using a vocal microphone or by means of filtering. Direct injection recording of the electrical signal from the pick-up often causes problems with hum due to ground loops. Safety problems may be solved through the use of a direct injection box or specially designed microphone amplifiers without a ground connection to the mixing console. Often the direct signal is mixed with that of a microphone.

The electric bass, like the electric guitar, may be recorded with a microphone. If the electric bass does not appear musically up front, it may be fed directly to the console from a pickup. Of course both methods may also be combined. This is often the case for a microphone pick-up when problems arise with feedback.



A. Classical or acoustic guitar recording



B. Amplifier microphone placement

chorus:	voice doubling; subjective intensity increase.
compression sustainer:	compressor which lengthens a tone without decreasing level.
distortion:	non-linear distortion with adjustable properties.
flanging:	time shifted signal overlay; varying delay time; vibrato effect.
noise gate:	shut-off for modulation pauses.
overdrive:	tube amplifier type distortion; e.g. increases with increasing level.
phasing:	similar to flanging: phase shifted signal overlay; frequency response of comb filter curve; also time variable.
spectrum:	adjustable boost in a steplessly adjustable frequency range.
touch wah:	automatically scanned filter with every tone.
wah-wah:	individually controlled scanned filter.

C. Guitar effects devices

The recording of keyboard instruments

Piano

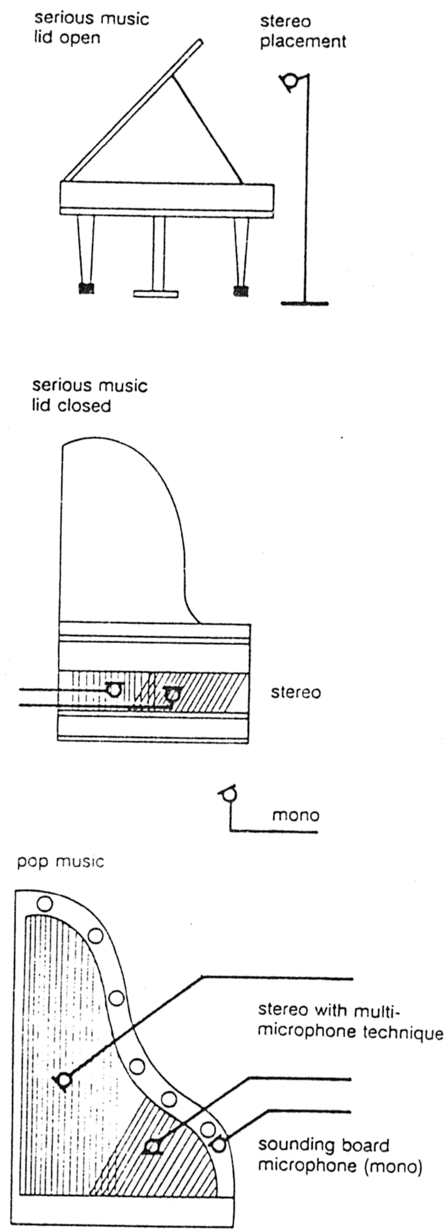
A piano is generally any stringed keyboard instrument with a hammer action. The proper name is actually pianoforte. For recording, one should always use a so-called grand piano (225 cm (7'4") or longer). For serious music recording, unless it only functions as a percussion instrument in the orchestra, the instrument is best recorded in stereo and the room tone is to be included in such a recording. This means a microphone spaced not too closely, and it may be done in MS, XY, AB or one of the other techniques. There is no single preferred method; however, the time-of-arrival and the mixed techniques appear to provide the best results. Many singers prefer a closed piano lid. In such cases a position to the right of the pianist or above the music stand provides the best presence, although disturbed to some extent by page turning noises. Basically a closed lid provides no advantages and must be avoided. A proper tonal balance is always possible with an open lid as well and complete removal of the lid is best of all. (The hinge pins come out of every grand piano to make removal simple). A closed lid always makes the grand piano sound muddy and dull, making it virtually impossible to use any satisfactory stereo recording technique (fig.A). For popular music it is recommended to place one microphone each above the low and the high register strings at a distance of about 10-20 cm (4"-8") and to use a pan-pot to obtain the proper stereo width. For monophonic recording, a suitable microphone position is over the high strings or over one of the front holes in the frame. Even though the sounding board, as viewed from below, is attached to the frame, a microphone underneath the piano yields a dull and tubby sound. The sounding board of a piano is quite thick (approx. 12 mm (0.5") and, therefore, does not vibrate very much while the strings give off their high frequency harmonics directly. For the upright piano, seldom used in recording, the lid must be opened for popular music and the individual microphones placed as shown (fig.B).

One unsolvable problem found in recording a piano and a single singing voice or instrument is the dual-roominess of such recordings. It always happens when two sound sources are of unequal loudness and are brought to equal loudness at the console. The louder sound source is then still represented more loudly in the reverberation signal. Therefore, the louder instrument always sounds further away and in a larger room, the softer one always nearer. The dual-roominess is also caused by the fact that the piano tone's decay has the same structure as the reverberation of the room, both spectrally and in duration. Thus the instrument's own reverberation masks the room reverberation and the roominess is only marginally audible. On the other hand the singing voice, for instance, can cause the roominess to be especially clearly audible; individual, impulse-like sounds generate single reflections which permit the roominess to come out clearly causing the singer to appear to be in a larger room than the piano.

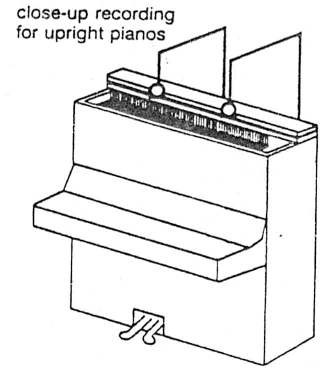
The excellence of the instrument condition is of paramount importance in recording (tuning, voicing). The piano make and, with it, the sound timbre of the instrument do play an important part in practice. Certain makes of piano are preferred by certain pianists and for certain kinds of compositions.

Other keyboard instruments

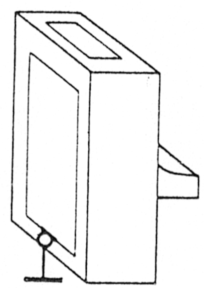
The celesta (fig.C), a keyboard instrument with the action of a piano but with metal rods instead of strings and individual wooden resonators has often been required in orchestras since the end of the 19th century. A microphone position at the back

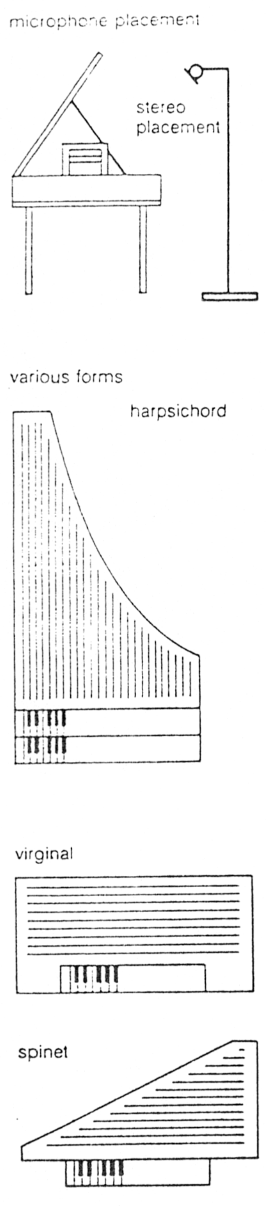


A. Grand piano recordings

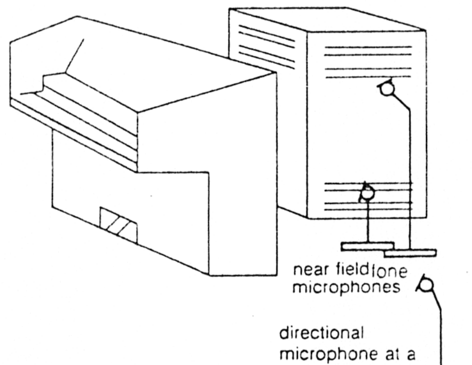


B. Upright piano recording

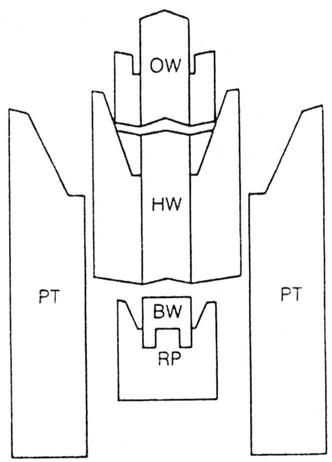

C. Close-up celesta recording



D. Harpsichord



E. Leslie speaker recording



F. Organ layout
(according to the division principle)

HW	Hauptwerk	(great)
OW	Oberwerk	(swell)
BW	Brustwerk	(choir)
RP	Rückpositiv	
PT	Pedalturm	(pedal tower)

of the instrument is preferred, in order to prevent the striking noise itself from becoming overbearing in a recording.

The harpsichord is a keyboard instrument whose action plucks the strings. It normally has the form of a grand piano, but smaller variations are also used. The triangular spinet (more rarely five or six sided) and the rectangular virginal (fig.D) are encountered more rarely. These instruments are now available in a more modern construction with heavier materials and a sounding board open at the bottom as in the piano, but of late there appears to be a trend back to traditional constructions with a closed sounding box. The microphone setup is largely analogous to that used for the piano (fig.D). The even loudness of the tones of these instruments is characteristic for them. This may only be altered through changes in registration, articulation and through the density of the musical composition as well as a comparably high subjective loudness level for the same electrical output levels as other instruments. The high loudness level is caused, in part, by the great sound density and a spectrum extending up to the highest frequencies. Therefore, in practice the harpsichord is often peaked at only 50% of full output level (-6 dB). By contrast, the clavichord is among the softest instruments in music altogether. Here the recommendation is to place a microphone very close over the tiny sounding board located at the right hand end of the instrument.

Keyboards

The keyboard instruments in popular music, aside from the piano, include the electronic (electronic organ, synthesizer, strings, etc.) or electro-mechanical instruments (electronic pianos, Hammond organ). The sound is given off through loudspeakers or is available for recording as an electrical signal. As in the electric guitar, a microphone may be placed directly in front of the loudspeaker although direct injection into the console without a microphone is preferred today (→p.120).

This is not possible when using a Leslie loudspeaker because the typical Leslie sound is created through rotating elements in front of the loudspeakers. The rotation produces a pitch vibrato due to the doppler effect which, if produced electronically, does not yield the same sound. A slow rotation causes the *cathedral effect,* a fast one the Leslie effect. At least two microphones must be set up, since the highs and lows are radiated from two separate loudspeaker systems (fig.E). Wind screens are required because of the air motion. However, it is better to place a highly directional microphone at a greater distance, since otherwise the intended pitch vibrato is superimposed onto the loudness vibrato.

Pipe organ

The large pipe organ such as a church or concert organ is the largest instrument considering outer dimensions and amount of material used. There are various schemes for the arrangement of the organ pipes. Following the lead of baroque instruments, groups of stops are combined into so-called chambers which are each assigned to a certain manual (fig.F). Both for optical and acoustical reasons, the low stops of the organ pedals are divided between two pedal towers, one left and one right, in such a way that adjacent half notes are distributed to both towers. The result is that the bass melody constantly jumps back and forth between the extremes of the organ chambers. In organs of the 19th century, the visible pipes are arranged entirely for optical/æsthetic considerations, often even using dummy pipes. The individual stops are not combined into sound balanced chambers. Between these two versions any conceivable combination of optical and acoustical pipe arrangement principles exist. For large church organs the recording should impart a feeling of spaciousness, This requires greater microphone spacing. There are no preferred recording techniques.

The recording of speech

Playback loudness and frequency response

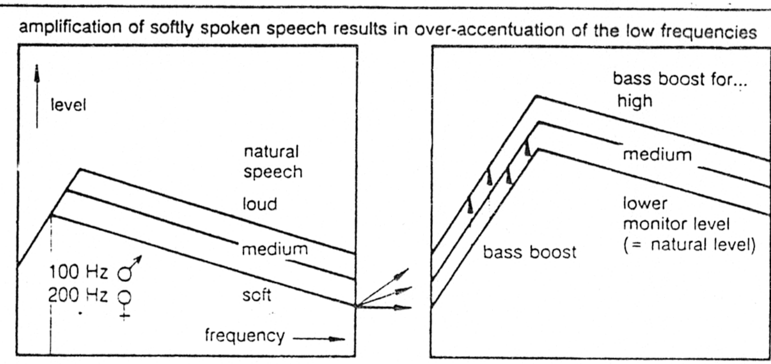
The connection between playback loudness, natural loudness and coloration exists in recording generally, but is particularly perceived in speech recording because the human voice is among man's best known sounds. The loudness of those voice components below about 100 Hz for men, and 200 Hz for women is relatively independent of the speaking volume (→p.60), and is therefore mainly determined by the distance from the announcer. In every acoustic reproduction in which the playback loudness level deviates from the natural loudness at the microphone location, there results an unnatural sounding reproduction of the lows; at unnaturally high levels the voice drones, because the lows are over-accentuated vis-a-vis the highs. For unnaturally soft reproduction the sound becomes flat because the lows are missing (fig.A).

Announcer recording

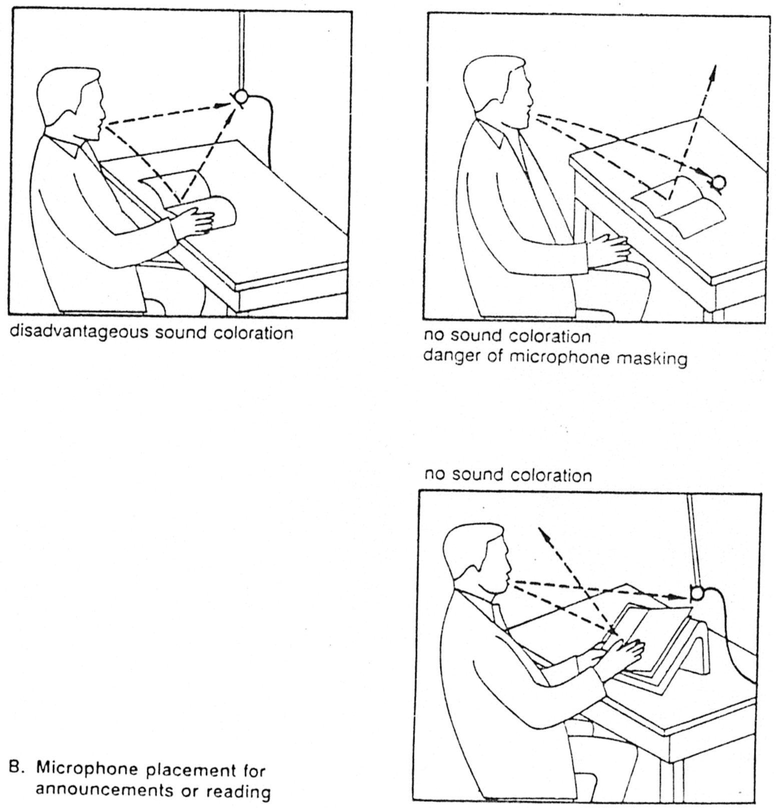
The sound pressure level for normal speech at a distance of about 60 cm (24") from the announcer is about 60 dB which increases by about 4 dB to 64 dB when this distance is cut in half. For loud speech, this level increases by about 6 dB. As a result, in a properly treated broadcast studio, the unweighted signal-to-noise ratio to the general studio and microphone noise amounts to about 50 dB. It is the self noise level of the microphone which is the determining one. Since the studio and microphone noises lie above the tape noise, short pauses in speech recordings must yield mostly studio ambience. Therefore, it is recommended that a recording be made of the studio atmosphere, consisting of studio, microphone and tape noise in case pauses have to be added to the recording at a later time.

For relatively close microphone spacing – below 30 to 50 cm (12" to 20") – the proximity effect (→p.70) and its noticeable low frequency boost provides a certain boominess to the sound. For such cases there are microphones with switchable bass roll-off, or microphones with a fixed low frequency droop, i.e., so-called vocal microphones (→p.74). For the 60 cm (24") spacing particularly common in studios, this effect does not play any significant role. Often the lavaliere microphone (→p.74) is used for speech recording and is worn on the chest. In spite of its strange position, no adverse response results. Much more annoying at close microphone distance is the popping caused by the explosive sounds of the speaker. A so-called pop screen is the solution. Disturbing sound colorations result when the microphone records reflections from the desk or manuscript along with the direct sound; a choice of the proper physical arrangement will help to avoid this (fig.B). Such sound colorations become disturbing when the comb filter curve (→p.107) resulting from these reflections shifts as a result of motion of the speech source. By contrast to dramatic studios, purely announce studios don't have a minimum size. Proper selection of the speaker's and microphone's location may provide room acoustics which will fulfill every acoustical need. The reverberation time is generally 0.2 s to 0.3 s; first reflections should be suppressed as much as possible.

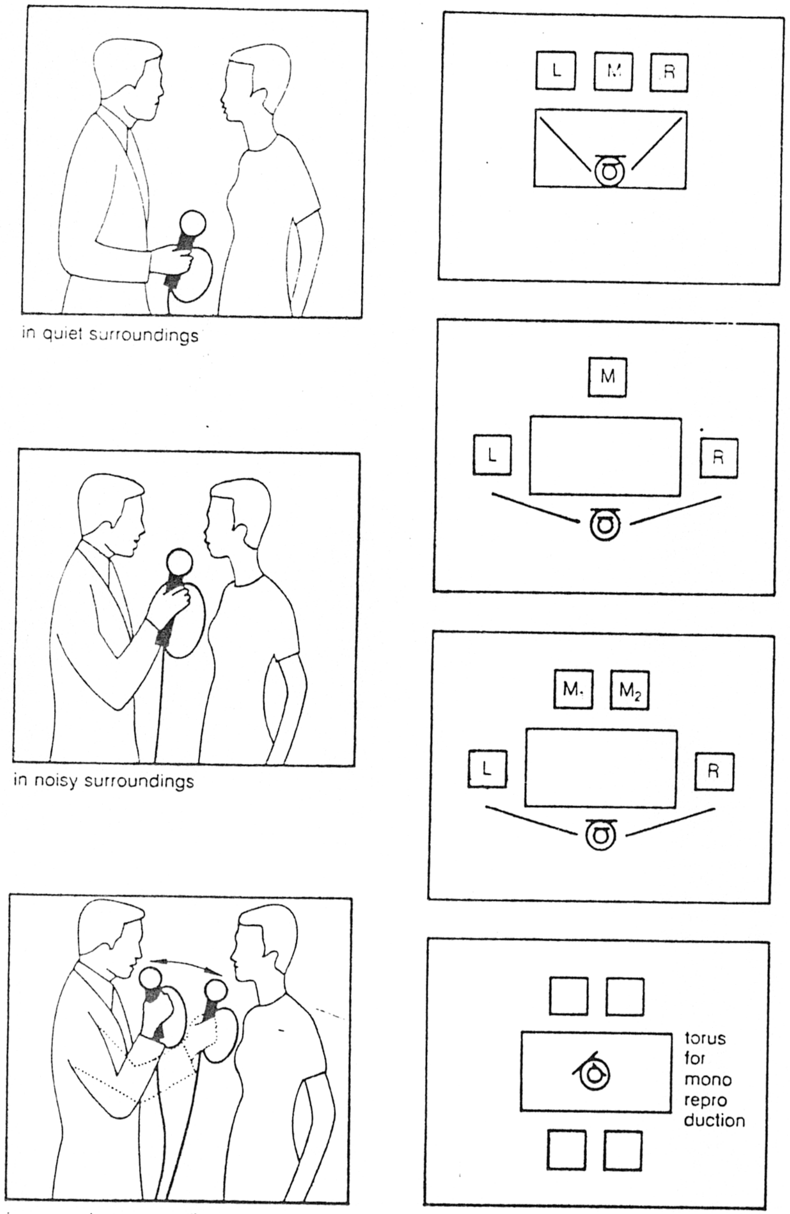
Several demands are made on the announcer – some from the engineering side, some from the problems of the medium. Strongly accentuated words at sentence beginnings will result in a reduced average level, thus to a reduction in intelligibility, and will have a negative effect on the loudness balance with other announcers or music. Pauses between parts of the text, for instance, should be shorter than for a lecture in front of an audience. Aside from this, recordings require greater discipline



A. Schematic presentation of the frequency response alteration resulting from "unnaturally" loud monitoring



B. Microphone placement for announcements or reading



C. Microphone placement for an interview

with regard to noises such as script page turning. Because of the close microphone spacing which acts as the surrogate for the listener when compared to the live situation, care must be taken to suppress noises such as those caused by saliva, etc.

Interviews, eye witness accounts

A suitable interview microphone is principally selected for its directional characteristic. An omni-directional microphone is suitable if one wishes to transmit the acoustic atmosphere along with the speech. Omni microphones are also less susceptible to wind, pop and finger noise than directional ones. The cardioid is well suited for situations in which unwanted noises are to be suppressed and only the interviewer or hissubject is to be recorded. The background noise determines how the microphone is to be held (fig.C). A figure-8 theoretically attenuates background noise just as well as a cardioid and is, therefore, well suited for the recording of two speakers, but must then be held at mouth level. Due to the figure-8's greater pattern integrity, it will attenuate especially low frequency background sounds best of all the directional patterns.

A wind and pop screen is always recommended. For microphones spaced closer than 30 cm (12"), one should use a close talking microphone (see above). However such a microphone falsifies the surrounding atmosphere. For very close talking, the microphone's membrane should not be addressed frontally, but rather at an angle to avoid popping noise and overload. Directional microphones are rather sensitive to mechanical noise interference. Therefore, it is important to avoid scraping or rubbing sounds against the microphone or even the cable.

Round table discussions

The same factors that are true for single announcers, such as sound pressure level, background noise, proximity effect and sound coloration apply to round table discussions as well.

There are two possible microphone setups. In the first each of the speakers gets a microphone according to the multi-microphone method (→p.94). The directionality for stereo then must be assigned through the use of pan-pots on the console. This method offers the possibility of opening a particular microphone only when it is needed. This function may also be fulfilled by a so-called noise gate. To prevent the appearance of an acoustic *hole* during pauses in the conversation, it is best to set up a microphone solely for room-tone. For qualitative reasons this method is most suitable to mono recordings. A better impressiön of the acoustical atmosphere in the discussion room may be provided by two stereo microphones at a somewhat greater spacing from the sound sources. The discussion partners are then best arranged on an arc of 270°, while in XY technique, two cardioid patterns are each rotated 45° from the center line (fig.D). A setup with two microphones back-to-back is not recommended.

A suitable setup for mono recordings is a stereo microphone with figure-8 patterns rotated at 90° to one another and combined through a 90° filter. This forms the directional characteristic of a rotating figure-8 (torus or *doughnut*), i.e in the horizontal plane equally sensitive all around but highly discriminating against diffuse sound from above or below.

The recording of vocal soloists and chorus

Vocal soloists (popular music)

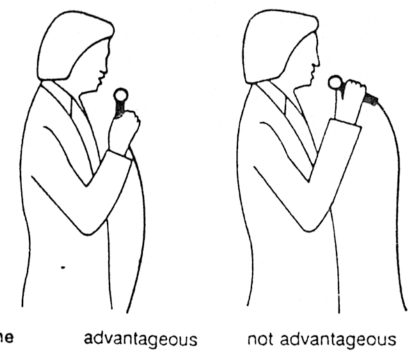
In this form of music the microphone-to-sound-source distance is basically very small. One often uses a hand held microphone to give the singer freedom of movement. The openings at the back of the microphone are not to be obstructed by the hands because this transforms any directional characteristic into an omni-directional pattern. If the vocalist also plays guitar or another instrument, then a microphone stand or a lapel microphone connected with a wireless system are needed. The close microphone spacing means high sound pressure levels, which, especially for explosive sounds (b, p, d, t), leads to popping and for sibilant ones (s, sh, z) to a scratchy sound. Both are caused by overload. The danger of such unwanted sounds largely may be compensated if one does not address the membrane head-on but rather from the side (fig.A). A pop or wind screen (or blimp) over the microphone also helps to reduce such effects markedly. It is indispensible and is built right into those microphones specifically intended for vocalist use. There are both dynamic and capacitor microphones available as vocalist microphones that meet all of the requirements for a close-talking microphone.

High quality dynamic microphones are often used in close talking applications. They are rather overload proof on the one hand, on the other very rugged. While directional microphones are normally used, one should also consider omni ones. The advantages of using a pressure transducer are lower pop sensitivity due to a more tightly stretched membrane, lower mechanical interference (e.g. from finger rubbing), and no proximity effect. By contrast, the problem of sensitivity to sounds from all around appears not to be too great a disadvantage, since high levels are produced from the very close spacing. For soft voiced singers, however, the danger of feedback from sound reenforcement systems may be too great. All directional microphones display an increasing bass boost with decreasing distance, resulting from the proximity effect (→p.73). For this reason, the microphones intended for vocalist use are either equipped with a bass roll-off switch or with a fixed low end roll-off filter (→p.74).

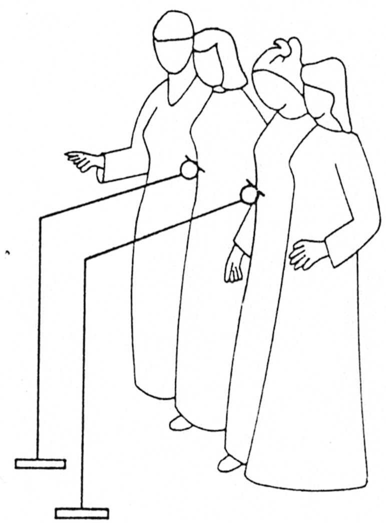
For the recording of smaller vocal groups or background choruses, one can use individual microphones under the same consideration as for individual soloists. The advantage lies in the ease of balancing the individual voices. Since one usually works with sheet music in studio recordings, it is recommended that music stands be used to facilitate page turning. For an internally well balanced group, one microphone may be used for two singers (fig.B).

Vocal soloists (serious music)

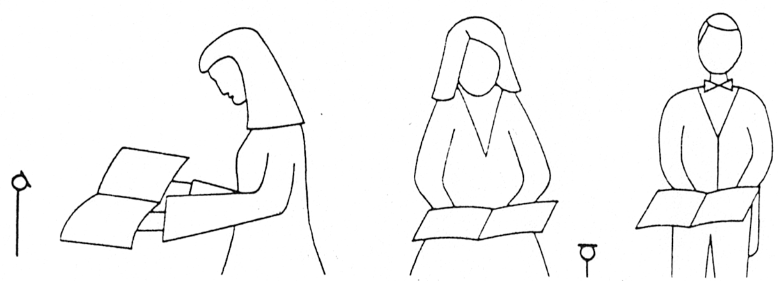
Hand microphones are never used for such singers. A microphone distance of 1 to 2 m (39" to 78") is preferred. Care must be taken in the microphone setup so that no sheet music is placed in the sound path between the singer and the microphone. Furthermore, attention will have to be paid to the concert situation in live recording to make sure that the microphone does not cover the soloist's face when viewed from the audience. A microphone height at about the level of the sheet music is preferred (fig.C), a height which avoids disturbing reflections from the sheet music itself. When there are several soloists, they may sing into one microphone in pairs (fig.D). Serious music soloists, especially sopranos, generate a considerable sound pressure level, so that the dynamic range may get to be extensive – greater than for most instruments (→p.63).



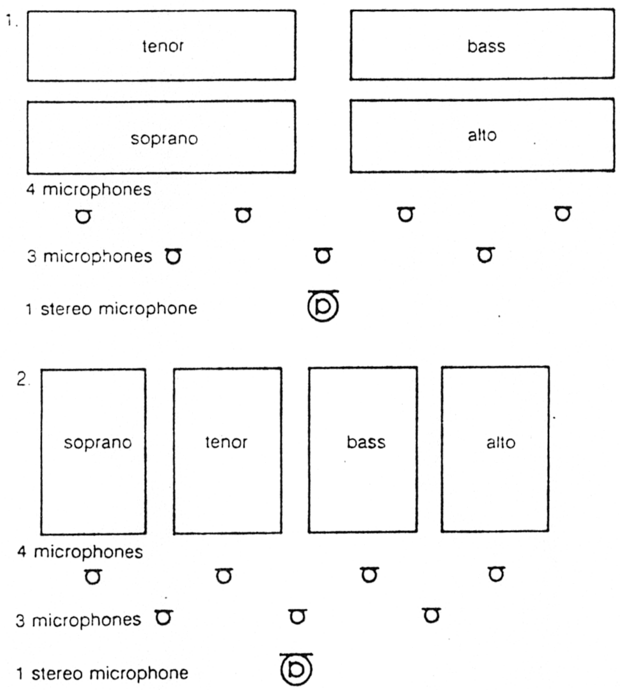
A. Hand held microphone for singers



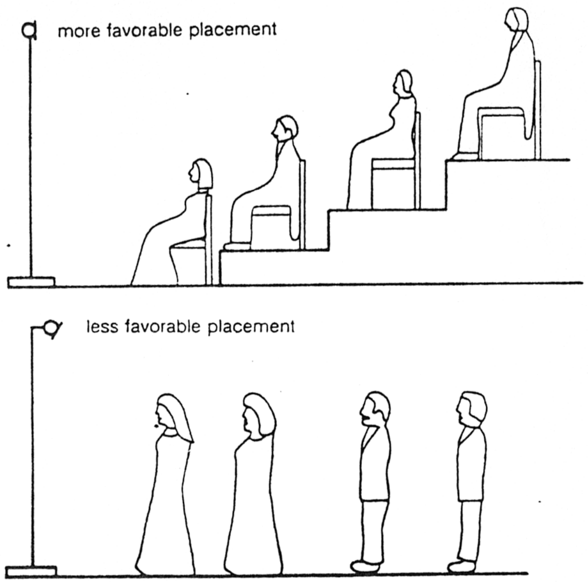
B. Singing group



C. Microphone for vocal soloists (serious music)



D. Placement of choral sections and microphones



E. Arrangement of choral singers

Chorus

The arrangement of voices (→p.60) within a mixed chorus usually follows the example shown in fig.D.1. It is advantageous with such a setup that those singers from each section who stand near the center have good acoustic contact to all other sections. In the performance this leads to a homogeneity of the choral sound. At the same time, however, it also leads to a certain scrimming effect as for instance of the counter-pontile structures. The arrangement according to fig.D.2 provides the listener with the ability to differentiate acoustically between the individual voices. This leads to increased transparency but, at the same time, makes it difficult for the voices to sing together. This arrangement follows the same criteria of sound symmetry which also produced the German orchestra seating. An arrangement according to the American seating is also possible: soprano - alto - tenor - bass. The singers should be placed on risers to permit the free radiation of their voices towards the microphone. Even a high microphone position cannot replace the riser method (fig.E).

There are various makeups for a chorus or choral composition:

Description of the chorus	Membership	Number of voices	Listing of voices from top to bottom	Special groups; remarks
Mixed chorus	Women (children) and men	Usually 4, rarely 5, 6, or 8	Soprano, alto tenor,bass; if more than 4 voices: soprano I and II	chamber chorus (small mixed choir) double chorus, (2 mixed choirs), a cappella chorus (w/o instrumental accompaniment)
Concert chorus				Usually amateur singers, secular and church music with orchestra;
Church choir				Lay singers; church music;
Opera chorus				professional singers; used in opera performance
Women's chorus	only female	usually 3 voices	soprano I, II and alto	
Men's chorus	only male	usually 4	tenor I, II and bass I and II	
Boys choir	only boys voices; often as mixed chorus with men's voices	1 to 3 4 to 8	soprano I and II and alto as in mixed chorus	

A stereo microphone is only then suitable, if the chorus does *not* stand behind the orchestra. In such an arrangement, the orchestra would sound too loud through the choral microphone, and would prevent proper orchestra/chorus balance. Such orchestra leakage would also tend to veil the orchestra sound. Three or four single microphones would be preferable in such a case, and would allow the balancing of the various voice groups.

Æsthetic principles in musical recording

As with the æsthetics of all art forms, the æsthetics of sound are forever in a state of flux. This field, like all others, experiences short-lived trends or developments connected with a particular individual. Nevertheless, it should be possible to formulate some classical æsthetic principles in stereo recording. The treatment of the æsthetics of sound gain in importance the more complex, the more spatially extensive a composition or performance is. In a recording of a singer with guitar it is relatively unimportant whether the singer is imaged to the right or to the left of the guitar, or whether the guitar is a bit closer or more distant than the singer. In a recording of a large work with vocal soloists, solo instruments, orchestra and chorus, on the other hand, the æsthetics of the sound take on the dimensions of an important artistic question, which, as with the performance of singers and musicians, is part and parcel of the work's interpretation. Æsthetic decisions about the sound usually start with relatively small ensembles. They may not be made independently of the performance necessities and, therefore, often take on the character of a compromise. Of course, it is a fact that the traditional placement schemes were obviously all developed under consideration of the sound æsthetic needs. A good example of this may be found in the German orchestra seating (→p.34). It may well be that the current trend towards greater precision in ensemble playing is often of greater importance, and this, in turn, leads to the American seating.

In spite of all the changing aspects of sound æsthetic judgments, there are two æsthetic principles applicable to stereo recording which are immutable: symmetry and clarity. The transfer of these basic principles to the individual sound and space dimensions of a stereo recording may be accomplished if the following guide lines are followed.

Distribution of sound sources along the stereo horizon: The sound sources are symmetrically distributed along the stereo horizon, according to their pitch: high pitch (left) - low pitch (center) - high pitch (right). This distribution is to be preferred to the low – high – low arrangement because it is the higher frequency components which more clearly define the flanks of the stereo image and because the low pitch instruments do not let the problems of the phantom sound sources (→p.78) appear as clearly. But the low pitched instruments also belong in the center functionally because they form the common harmonic fundamental of all the instruments. In this particular sense, it is the German seating which is preferred. There also is a great advantage for the cutting of stereophonic records, since low frequencies waste the least space when centered. Large works, for instance for vocal soloists, orchestra and chorus, have several sound layers in the spacial depth, e.g., the vocal soloists in front slightly behind them the orchestra and, behind it, the chorus. If the sound sources are well distributed among the various distance layers, it clearly increases the clarity and transparency of a stereo image. For this purpose there exist the following methods for assigning the sound sources to two spacial layers, conforming to their pitch:

low – high – low
high – low – high

high – low – high
low – high – low

high – mid – low
low – mid – high

low – mid – high
high – mid – low

For three perspective levels, the elements which are possible for two levels, may be combined in many ways but at all times keeping the basic principles of symmetry and clarity in mind. Here are two examples:

high – low – high
low – high – low
high – low – high

high – mid – low
high – low – high
low – mid – high

A single solo instrument or singer is always placed at the center of the horizon. The fact that in practice such sound-æsthetic considerations are often in juxtaposition to performance based habits and demands in no way limits their validity. In many situations it should be easy at least to heighten the sound symmetry and clarity through placement of the vocal soloists.

Width of the stereo horizon: The stereo width should not run counter to the room perspective. Wide sound areas should be reproduced as wide as possible; narrow ones, as for instance two instruments, are reproduced with decreasing width for increasing distance of the sound source. A wide image horizon does not contradict the spatial perspective of the listener only for closely spaced microphones. The room tone always fills the entire horizon width independently of the reproduction width of the various sound source.

Depth perspective: The smallest acoustically reproducible depth with loudspeaker reproduction is the spacing of the loudspeakers; the largest possible acoustically reproducible distance altogether lies between 10 and 20 m (33 ft. and 66 ft)(→p.78). Since the distance perception ability is not as well developed as the directional perception ability, it only is possible to differentiate between a very few distance planes from smallest to largest reproducible distance. A well defined depth differentiation in no way matches the experience of natural hearing, at least in the music field, but rather primarily results as an acoustical perspective from the position of the main microphone, and from the position of the conductor as well. The depth differentiation offers the possibility of differentiating the sound space, a possibility which should be utilized fully to compensate for the narrowing of the roominess which is unavoidable in loudspeaker reproduction. Depth differentiation, at the same time, also means differentiation of meaning. From our daily aural experience we learn that that which is nearby is that which matters or even threatens; in short that which is more important than that which is distant. The depth differentiation realizable in a recording may be that much more defined the larger the playing group and the larger the perceived space. Acoustically it is not easy to generate depth perception (→p.104).

Semi-classics, popular music, folk music, jazz: The previously mentioned sound æsthetic criteria are just as applicable to these types of music as they are to serious music. The distribution of sound sources according to their pitch is perhaps even more important here because the localizability on the stereo horizon, resulting from the primarily multi-microphone and support microphone techniques, is even better. The definition of the sound planes must be looked at somewhat differently. For example, the rhythm group with drum set or rhythm and bass guitar or, perhaps even the drum set alone, represents a sound plane. It is logical to place the bass drum in the center, the cymbals and hi-hat are positioned at the extremes while the tom-toms are arranged in ascending pitch order from left to right on the stereo horizon. The spacial depth perspectives of this type of music does not have the same meaning as it does for serious music if for no other reason but the fact that the ensembles are much smaller.