Chapter 3
Tonal Characteristics of Musical Instruments

3.1 Brass Instruments

3.1.1 The French Horn

3.1.1.1 Sound Spectra

It is typical for brass instruments to exhibit spectra which can be divided into two groups. In the upper register the fundamental is strongest, while for the lower positions a formant-like maximum is present. As noted in the chromatic representation in Fig. 2.4, the fundamental dominates from C_4 on upward in French Horns, while higher partials decrease in amplitude rather steadily. Below C_4 the maximum is initially relocated to the octave partial and then maintains its frequency position, so that in the lowest registers the 4th and 5th partials receive most of the energy. As a result, the main formant, typical for the French Horn, develops, which is located at approximately 340 Hz (Meyer, 1967b). It falls into the region of the vowel color "u (oo)" which is responsible for the round and sonorous sound of the Horn.

Below this maximum, the amplitude for the low registers drops rapidly with a slope of 12 dB/Octave. The lowest frequency possible corresponds to the note B_1 with approximately 62 Hz. The fundamental is roughly 25 dB weaker here than the strongest partial. This shows that the low frequencies only play a subordinate role in determining the tone quality of the Horn. Above the maximum, the amplitudes also decrease, however, there are several additional formants present which influence the tonal picture. Their frequency locations are represented schematically in Fig. 3.1. Below the Vowels are the results for the commonly used German double horn. The already mentioned main formant at 340 Hz is followed by the first ancillary formant at 750 Hz, also in the range of the vowel color "a (ah)," and further formants near 1,225 Hz, 2,000 Hz, and also (using the F-Horn) still near 3,500 Hz. This series of ancillary formants brightens the overall tone, so that the tonal character is not as dark as for a sung or spoken "u (oo)." In this, the higher frequency partials gain in importance with increasing overall volume, while only the lower formants contribute while playing softly.
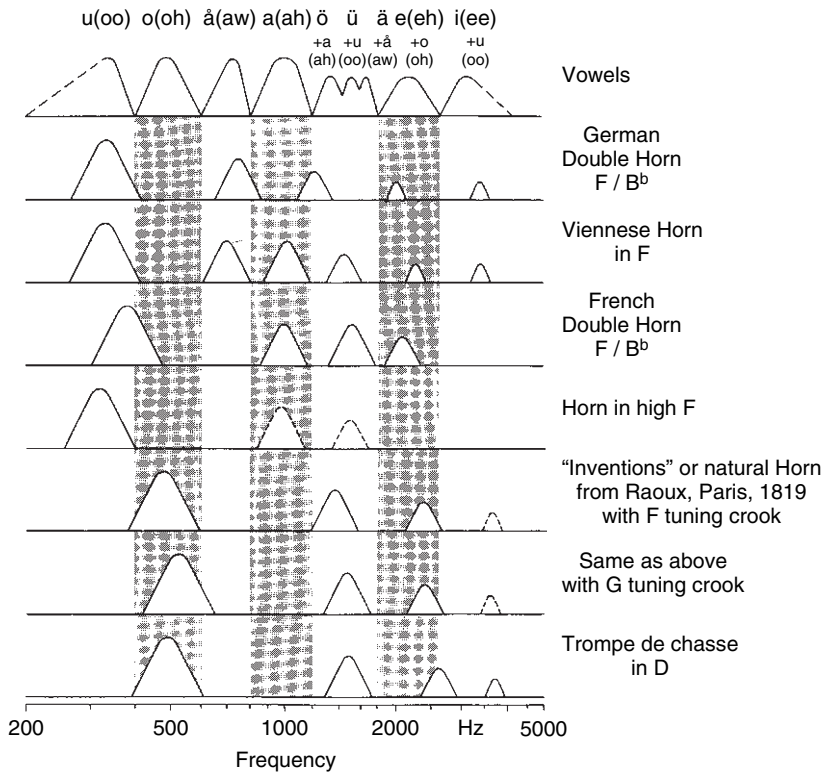
Fig. 3.1 Schematic representation of formant position for various horns

The intensity of the lower partials, and thus the fundamental tonal characteristics, however, also depend on performance techniques: When the mouthpiece is pressed relatively firmly against the lips, their vibrating portion is more clearly delineated, so that strong lip vibrations are possible without time dependent irregularities. Furthermore, instrument resonances are not damped as much as is the case for a low pressure contact. Consequently, for firm pressure, a fuller sound is produced, which carries better. In contrast the advantage of a low pressure attack lies in greater ease of playing in the upper register, which, however, is purchased with a thinner timbre lacking somewhat in substance. This also constitutes the typical difference between the European and American "schools" of brass performance technique.

The frequency range of the spectra extends upward to 1,500 Hz for the lower register at a dynamic level of mf. It increases to about 5,000 Hz for the highest notes. The overtone content of the F-Horn is somewhat larger than for the B^b-Horn, thus the F-Horn has a somewhat more pronounced tone color, particularly in the mid range. The valves also influence the higher frequency partials slightly: For fingerings with several valves the overtone content decreases so that the sound loses

brilliance and becomes dull. As a result the player receives opportunities for tonal variations in addition to intonation corrections.

Noise contributions are especially weak for brass instruments, so that their influence on the sound picture of the horn is negligible. Only in the lowest octave of the tonal range could minor noise components, up to approximately 3,000 Hz, be noticed above the partials. This can be described as a hissing with an "i(ee)" like tone color in relation to the corresponding formant.

3.1.1.2 Dynamics

The strong dependence of overtone content on dynamics in the upper register of the horn has already been mentioned in the previous chapter (see Fig. 2.7) as an example for the note f_4. As a generalization, the power spectra can be described in terms of a level difference of the order of 10 dB for *ff* and 50 dB for *pp* when comparing the strongest partial with the 3,000 Hz component. For high frequencies a level drop of only 5 dB/octave for *ff*, and 15 dB/octave for *pp* is associated with this. Furthermore, the influence on the timbre connected with this effect is enhanced by the shift of the main formant toward higher frequencies with increasing loudness, thus brightening the main tone color from "u(oo)" for *pp* and *mf* to "o(oh)" or possibly "å(aw)."

In the low register as well, a similar influence of loudness on tone color is noted. Thus in the second octave, the power spectra show a level difference of 20 dB for *ff* and 50 dB for *pp* between the main formants and 3,000 Hz, which is associated with a level drop at high frequencies of 11 dB for *ff* and 15 dB for *pp*. However, by reason of the narrower partial positioning, there is still a group of approximately 8–10 harmonics present. The amplitude differences between the respectively strongest partial for *ff* and *pp* lie at roughly 20 dB. They are thus slightly larger than at higher registers. In addition, a shift of the dynamic dependent amplitude maximum is also noted in the low register. For increasing loudness the vowel color brightens, though this effect is not as pronounced as in higher registers. The dynamic and modulation possibility of the natural horn sound thus rests both on a change in tone color in the region of the lower partials and on the possibility of large amplitude changes for the higher frequency sound contributions.

For rapidly performed scales, the horn reaches power levels of 107 dB for *ff*, while for *pp* 86 dB are produced. For individual notes, the lowest values of 65 dB are reached. These, as well as the upper limits of playable dynamics are shifted over the tonal range of the instrument toward greater loudness, where, for a high range *ff*, 117 dB certainly is possible. On the whole, a practically realizable dynamic range of 35–40 dB can be considered as typical for the horn. However, for the highest notes a genuine *pp* can hardly be expected. (Meyer, 1990).

A sound power level of 102 dB can be given as a characteristic value for the *forte*-sound. The influence of dynamics on tone color expresses itself as a level change at 3,000 Hz. This increase ranges from 1.5 dB in the low register, to 3 dB in the higher regions – accompanied by a simultaneous change in the level of the strongest partials by 1 dB.

3.1.1.3 Time Structure

The initial transient for a tongued tone of a horn is characterized by a short pre-cursor impulse, well known from other brass instruments (see Fig. 3.5). The duration of this impulse, which contains primarily harmonic partials below 1,000 Hz, lies in the order of 20 ms. It begins, depending on sharpness of attack, between 10 and 30 ms after beginning to excite vibrations (Melka, 1970). Several such impulses can follow, which gives the onset of the tone a character of a rolled "r" which naturally is esthetically undesirable. As the attack is softened, the pre-cursor impulse diminishes in importance in the tonal picture. An excessively strong pre-cursor impulse with slow development of the fundamental, on the other hand, generates the notorious "blare."

The duration of the starting transient for tongued tones is shortest in the middle and high register, above F_3 it amounts to approximately 30–40 ms, in the low registers it rises to values between 40 and 80 ms. (Melka, 1970). These values are confirmed for tongued notes in the sonograms of Fig. 7.26, which will receive closer attention in the context of room acoustical effects. For soft attacks, the initial transient can last longer than 1/10 s, which does not include consideration of dynamic development for long notes.

Inasmuch as the air column in the instrument is capable of storing only small amounts of energy, the decay time of brass instrument notes is relatively short. Even the energy stored in the wall vibrations does not increase the decay time significantly. The horn typically has decay times around 150 ms.

When going from one note to another it is important to note whether the transition is accomplished by the lips or a valve. In the former case a continuous frequency transition is observed which serves to smooth out the connection, while the activation of the valve leads to a sudden frequency change in the resonator and a break in the vibrations. This makes the transition harder or more pronounced (see Fig. 3.3), so that either technique has tonal advantages depending on the musical context. In fact lip transitions are naturally only possible for certain tone sequences, furthermore, valve connections provide the player with an increased sense of security for the attack. A typical example for this is the triad motif in the large E major aria of Leonore in the first act of the Opera "Fidelio." This passage was originally written for natural horns, nowadays it is, however, mostly performed with valve transitions.

The impact of using a *vibrato* in horn playing is largely determined by stylistic considerations. When a vibrato is played, it appears mostly through amplitude modulation of higher partials of like phase, that is, through a pulsating tone color modulation, as shown in Fig. 3.6 for the lip *vibrato* of a trombone. The depth of modulation can reach approximately 10 dB for the higher frequencies. It thus contributes to the ability to notice the horn in an ensemble.

3.1.1.4 Special Playing Techniques

Normally the right hand of the player lies only loosely in the bell of the horn, causing a certain damping of the higher partials. The possibility of using the hand to close the bore almost entirely has been developed as a special effect to change the
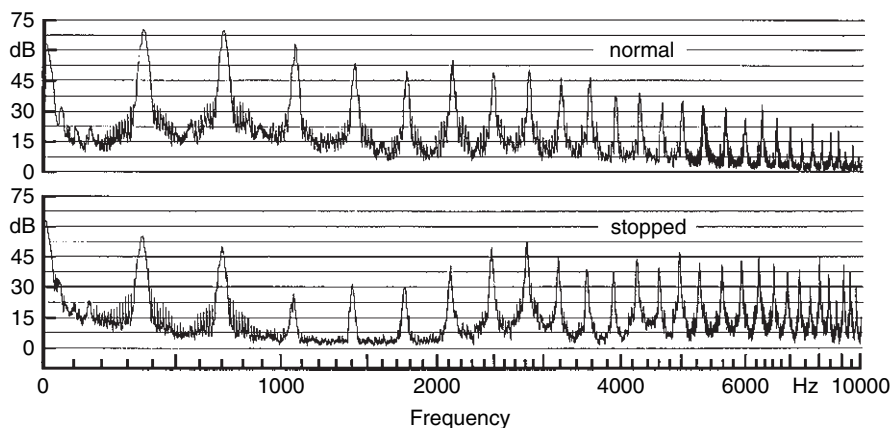
Fig. 3.2 Tonal Spectra of a Horn for the note F_4

tonal character in a fundamental way. This is particularly pronounced in playing *forte*; thus, a "stopped *sforzato*" is often indicated, the tonal effect of which is "metallic brittle and rough" (Kunitz, 1961). A typical example for this is the final chord in the Beckmesser-motif of the "Meistersinger" (score example 3).

As shown in the analyses of Fig. 3.2, the stopped sound lacks several essential components, while others are strongly developed. The gap from the 3rd to the 5th partials is noticeable, which covers the region of the a(ah)-formant, causing the pressed tone, lacking in strength. On the other hand, the metallic timbre is strongly emphasized by the maximum at 3,000 Hz and the strong partials above 10,000 Hz.

A comparison of stopped tones at different pitches shows that these frequency locations for the typical maxima and minima are always maintained. As a result, the *ff* – sound of a stopped horn becomes even less substantive in the upper registers. However, for a *mezzoforte,* stopped tones keep their strongest sound components in the usual location. Yet even here, a weakening in the region of the a-formant and an increase around 3,000 Hz leads to a tonal change in the direction of a metallic timbre.

A different tonal effect is caused by playing the horn with upward pointing bell and without a damping hand. The best known examples for this are found in the symphonies of G. Mahler. However, intonation suffers slightly with this playing technique, furthermore, the tone becomes hard and "coarse" as a result of the stronger contributions of the higher frequency components in the stationary and transient part of the sound. For these reason the technique is avoided by many players and conductors (Kunitz, 1961).

3.1.1.5 Horns of Special Design

In the Vienna Philharmonic, even today, pure F-Horns are played. They are distinguished from the German double horn by their narrow bore, the so-called

Score example 3 R. Wagner, Die Meistersinger von Nürnberg, motif of "Beckmesser" (3rd Act, 4th scene)

"Vienna valves" (Stechbüchsen-Ventile) and a particularly shallow cone in the mouth piece. This design does maintain the position of the main formant near 340 Hz, that is, in the range of the vowel color "u(oo)," as seen in Fig. 3.1. However, the number and positions of secondary formants is noteworthy. This is the only group of horns which shows five secondary partials, where the first two are located in the vowel range "a(ah)" and "å(aw)." This double a(ah)-formant give a particular strength to the sound and effects a song-like character; in its spectral distribution it is also reminiscent of the violins of Guarneri del Gesù (Lottermoser and Meyer, 1962b), famous for their rich and powerful sound. The two following secondary formants are positioned slightly higher than in the German Horn, however, the high components in the Vienna school are damped slightly more using the hand, so that the sound in spite of its fullness displays a certain softness.

For valved note transitions, noticeable differences between the Vienna Horn and the German double horn also occur, as shown in Fig. 3.3. The rotary valve of the double horn provides several paths for the air stream, so that short duration turbulences occur. These are noted as crack-like transition noise, as recognized in the three-dimensional spectrum (Windholm and Sonneck, 1988). In contrast, the note transition with a pump valve progresses more smoothly; i.e., the Vienna Horn has the advantage of a softer transition and a better legato, where the advantage of the German Horn can be seen in the better articulated transition and the better staccato.
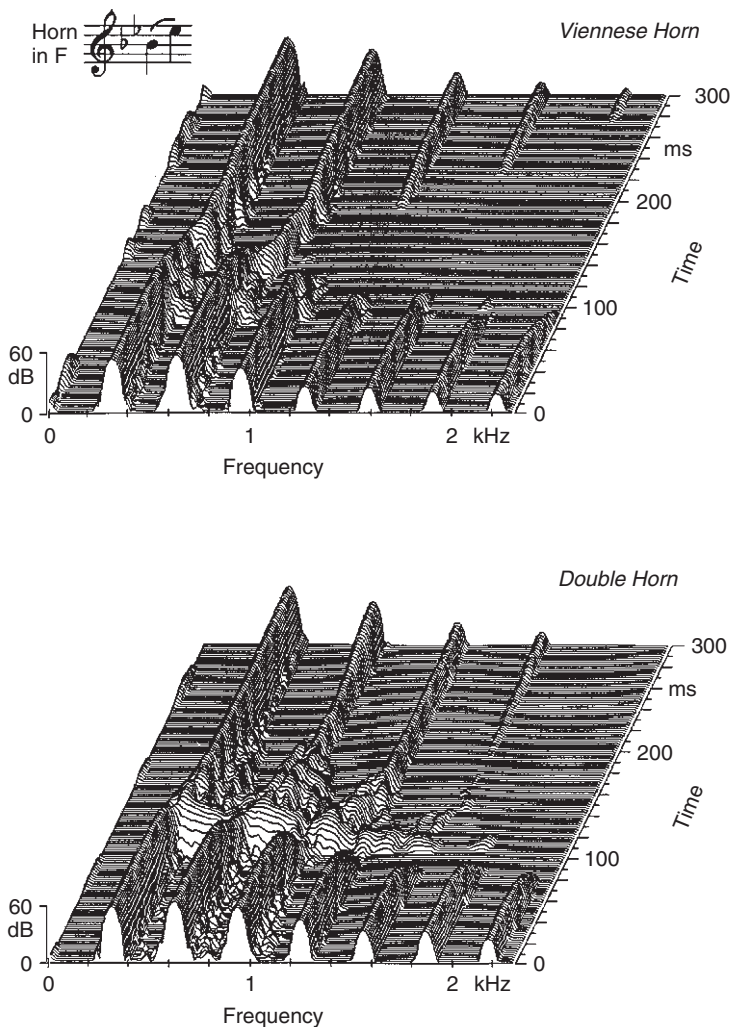
Fig. 3.3 Time development of a tone spectrum for a connected tone transition, represented for two horns of different construction (after Widholm and Sonneck, 1988)

Out of the tradition of horn virtuosity common in France, a double horn was developed, in use there today, which is distinguished by its slender tone and ease of attack. High intensities for upper frequencies create a relatively bright tonal character, somewhat reminiscent of a hunting horn. A change in the frequency location of the formants in comparison to other horns has a similar effect. The principal maximum is moved approximately to the boundary between the "u(oo)" and "o(oh)" regions, causing the basic vowel-like characteristic to become brighter. Two strong secondary formants are located in the region of the vowel "a(ah)," and also in the region of nasal components. It is precisely these strong

tone contributions around 1,500 Hz which form a particular characteristic of the typically French timbre. They also appear in bassoons in similar fashion. A further secondary formant in the frequency region of the vowel "e(eh)" brightens the tone of the French horn additionally.

Horns with high tuning are occasionally used, because they allow certainty of attack for high passages with greater reliability. However, this advantage is purchased with a reduction in tonal quality. For example, in a horn in high F, partials in the mid-register only reach up to 2,500 Hz, while in a normal F – horn they are present up to 4,000 Hz for the same notes at the same loudness. Connected with this is the lack of secondary formants, as clearly represented in Fig. 3.1. It is especially this circumstance which contributes significantly to the lack of color in the character of high F-horns, which can be described as dull and blunt. Only from C_5 on up can the high F-horn be considered to be equivalent in tone to the lower instruments, so that it appears, its use can only be justified for high passages.

3.1.1.6 Historic Horns

Double horns of today's standard design have been available since approximately 1900, when the development of valves around 1835 made it possible to construct chromatic instruments. The typical horn of the period from about 1755 to 1845, on the other hand, was the so called "inventions" or natural horn, for which the tuning of the natural tone series was adjusted by inserting different lengths of tuning crooks. This instrument was blown with the right hand in the mouth of the bell, as are today's horns, however it had a brighter tonal character.

The formant locations for an instrument from the beginning of the nineteenth century are shown in Fig. 3.1 as an example of a natural horn. The relatively high location of the main formant is particularly noticeable: For an F-tuning it is located at approximately 480 Hz, and moves to about 525 Hz for tuning in G. Additionally it should be noted that it is lowered to about 425 Hz for tuning in E^b. This means that these horns correspond to a brighter or darker "o(oh)" in their basic tone color, in contrast to an "u(oo)" for today's horns. The secondary formants of the natural horn also lie relatively high, and the nasal components are more prominent. As a whole, the overtone content is greater than in current horns, at *mf* the harmonic content of the vibrations of the lower natural tones reaches up to 3,000 Hz, and go above 5,000 Hz for the upper registers, where the notes are richer in overtone content for the longer horn, i.e., for the lower tuning. However, the differences in tonal characteristics for different tunings are not as pronounced as the contrast with today's horn. With respect to the tonal brilliance and the richness of the tonal picture, the natural horns in view of their timbre are not as far removed from the trumpet as today's instruments, nevertheless, the typical horn character is basically preserved.

The horn playing technique as employed in Baroque times with the corni de caccia can best be compared with today's French trompe de chasse, which is also blown without the damping hand in the bell mouth. Because of this technique, a tone rich in overtones is produced; the spectrum extends almost up to 10,000 Hz for

intermediate intensities, and even for *piano* it contains partials up to approximately 4,000 Hz. The main formant lies in the region of the vowel sound "o(oh)" – much like in the natural horn, the secondary formants also show frequency locations which compare to the natural horn. Since noise components are relatively strong for higher frequency tone contributions, the open hunting horns give a rough and distinct metallic impression. As a result of the richness in overtones, the trompe de chasse is brighter in timbre than the natural horn, so that it comes close to the trumpet, and particularly the bass trumpet.

3.1.2 The Trumpet

3.1.2.1 Sound Spectra

Among the instruments of the orchestra, the trumpet is one of the richest in harmonics. Already for a *mezzoforte* the harmonic tone contributions of the low and middle regions of the tonal range extend above 5,000 Hz. In the upper regions, the boundary of the spectrum is pushed to approximately 8,000 Hz (Mühle, 1965). This results in a radiant and brilliant tone, with the further characteristic, that the region of strongest partials lies in a relatively high frequency range. However, below this maximum, the spectrum drops at the relatively flat rate of 6 dB/octave.

For today's standard B^b trumpet the playing range begins with an E_3 (165 Hz) neglecting rarely used pedal tones. The main formant of the sound spectrum is located at about 1,200 Hz and is pushed up to about 1,500 Hz in the fifth octave. As noted in Fig. 3.4, this means an emphasis on the vowel "a(ah)" for the largest part of the range, responsible for the strong tone of the instrument. For the higher regions the nasal components become more apparent, without, however, removing the sparkle from the tone. In this, especially the secondary formants in the vowel regions "e(eh)" and "i(ee)" play an important role in the brightening of the sound. The prominence of these two groups of partials prevents an extreme sharpness which could arise from a tone so rich in overtones in the absence of formants.

The light and brilliant tonal effect is roughly uniform over the entire tonal range since the fundamental does not dominate the trumpet sound even while playing *mezzoforte,* except for the highest notes. The brilliance of the timbre is furthermore supported by the fact that the noise contributions are very weak, so that hearing impressions are hardly influenced at all.

3.1.2.2 Dynamics

With increasing loudness, the overtone content increases dramatically, so that for *ff,* tone contributions to the threshold of hearing are present. Under these circumstances the trumpet becomes the orchestral instrument richest in overtones. On the other hand, in the lower loudness regions the spectrum of high notes, similarly to the case of the horn, is reduced to a few partials, so that the timbre can become softer for example than that of the oboe.
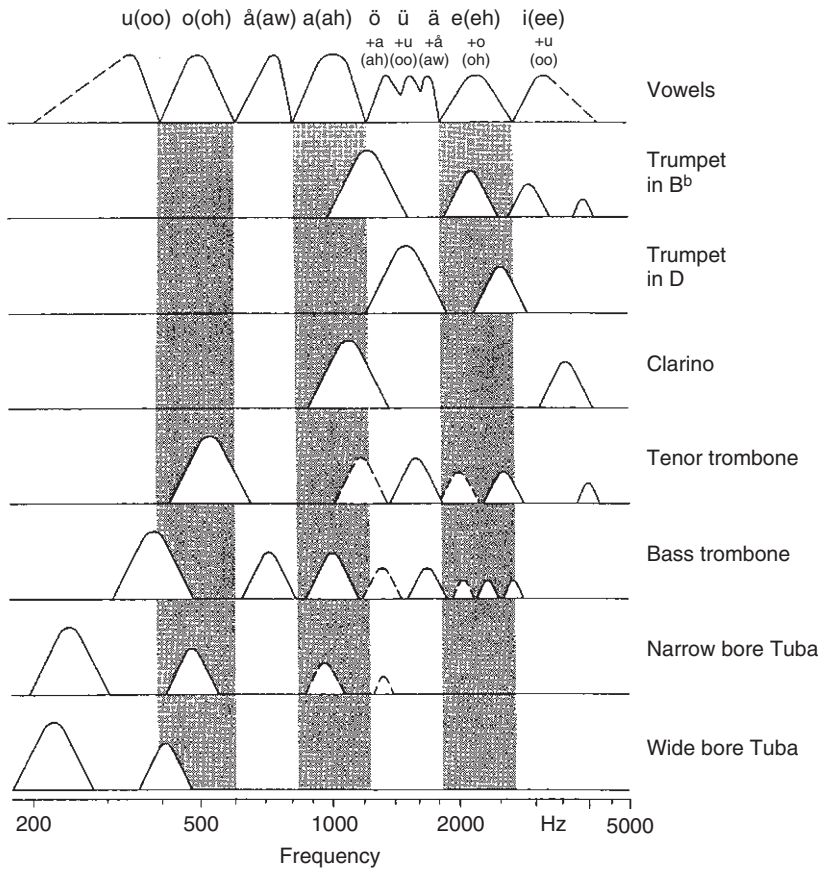
Fig. 3.4 Schematic representation of formant location for various brass instruments

In the power spectra, the 3,000 Hz level at *ff* lies by only 12 dB below the level of the strongest component, for *pp*, however, by around 40–50 dB. While the spectra for *ff* in the low register drop at 11–16 dB/octave by about 2,500 Hz, for *pp* this drop for the low notes already occurs around 1,600 Hz and for the higher notes around 2,000 Hz, and can reach a slope of 25–30 dB/octave, and for the highest notes even 50 dB/octave. The directional dependence of the sound radiation, however, is also very important for the tonal effectiveness of the high frequency contributions. This can lead to a significant intensity increase in the axial direction, when compared to the sound power levels averaged over all directions (see Sect. 4.2.1).

The dynamic range for rapidly played note sequences is characterized by a sound pressure level of 89 dB for *pp* and 104 dB for *ff*. For single notes the *pp* level may be dropped to as low as 78 dB in the low range and raised to 111 dB for high notes at *ff*. From this, a practical dynamic range of from 25 to 30 dB is obtained. The highest notes of the trumpet, however, lend themselves to less dynamic variation, which results from the relatively loud *pp* with a sound pressure level of 100 dB.

A level of 101 dB should be considered as a typical value for the sound pressure level of a trumpet at *forte*. The influence of dynamics on tone color is particularly extreme for the trumpet: For a variation of the strongest partials by 1 dB, the 3,000 Hz components of the low register change by 2.2 dB, while for the high notes, this results in a change by more than 4.5 dB (Meyer, 1990).

3.1.2.3   Time Structure

The initial sound of tongued notes is marked by an extraordinary incisiveness in a trumpet. Relevant for this is not only the comparatively short initial transient before the steady state is achieved, but also a very strong preliminary impulse. The initial transient for tongued notes in the fourth octave is accepted as from 25 to 30 ms. For higher frequencies this is shortened to values of below 20 ms (Luce and Clark, 1965; Melka, 1970). A first sharp peak in the amplitude representation appears, however, after 10–15 ms, depending on the sharpness of the attack. This preliminary impulse has a duration of only 5 ms in the trumpet, it is thus shorter than in the horn. It is already suggested in the amplitude development of the fundamental and the octave partial. The maximum amplitude occurs between 2,000 and 3,000 Hz, yet even for still higher components the pre-cursor impulse shows a higher intensity than the subsequent steady state. Accordingly, the precision of a trumpet staccato, is particularly achieved by the high frequency contributions. Even though the attack contains relatively few noise components, the attack can cause a crack-like impression by reason of the extremely rapid amplitude development in the pre-cursor impulse.

With a soft attack the trumpet tone develops very slowly. In the low register the initial transient can last for nearly 180 ms, in the middle register this is shortened only insignificantly to 150 ms (Melka, 1970). Only for very high tones is the difference between the duration of the initial transient for sharp and soft attacks no longer as large, and even a softly attacked tone develops within 40 ms, however the preliminary impulse is not as strong as for a staccato. Therefore, a tonal character which, on the whole, is soft, can be achieved in spite of the high overtone content.

3.1.2.4   Mutes

Use of a mute with the trumpet leads to a modification of the spectrum, this refers to the overtone content in general, as well as to the relocation of formant regions (Meyer, 1966c). The so-called normal (conical) mute strongly diminishes the intensity from the fundamental up to above 1,500 Hz. As a result, the "a(ah)" formant, which is so important for the open sound and for the tonality associated with the fundamental, is missing from the tonal picture. Accordingly, the timbre loses substance and it gives the impression of lacking strength. At the same time, the nasal contributions gain in importance, so that the muted trumpet sounds somewhat

squeaky. In addition, an increase in intensity above 4,000 Hz adds to the metallic character of the tone. A typical example for the use of this tonal effect is given by the excerpt from "Pictures at an Exhibition" orchestrated by M. Ravel, as shown in the score example 4. The color of the muted trumpet represents the verbosity of the imploring Jew Schmyle in conversation with the self-confident Goldenberg.

In contrast, the high frequency contributions are strongly reduced by the use of a cup mute, so that above 2,500 Hz, practically no partials appear (with the exception of notes in the highest playing register). Inasmuch as the intensity reduction is already effective from 1,000 Hz on, the formant is moved, depending on tone height, into the region of the vowel color "o(oh)" to "å(aw)" so that the tone is relatively round and without sparkle. The fast and precise initial transients, typical for a trumpet, however, remain essentially intact.

The so-called Wah-Wah mute finally makes it possible to influence the frequency location of its Helmholtz resonance by changing its position in the bell mouth. This causes time-dependent tuning shifts which are sensed as tone color transitions or flowing tone colors.

3.1.2.5 Trumpets in Other Keys

Inasmuch as composers in the classical era could only write their trumpet parts for natural instruments without valves, their works call for many different key trumpets. In today's orchestra trumpets with tuning other than B^b (with valves) are used only in special situations, and then particularly for very high parts, as they frequently occur, above all, in Baroque music.

The D trumpet, which in its tuning is located a major third above the normal B^b trumpet, shows in its tone picture a corresponding shift in its formant region to higher frequencies (Mühle, 1965). As recognized from Fig. 3.4, already for the largest portion of the tonal range the main formant moves to a position around 1,500 Hz, i.e., into the region of the nasal tone colors. As a result, the timbre loses strength, and shows an otherwise thinner effect. This tonal impression is further emphasized by the fact that the higher tone contributions, roughly above 2,500 Hz, are stronger than in the B^b trumpet. True, the difference in the fourth octave, on the average amounts to only to 2 dB, however, in the mid and upper registers it rises to more than 5 dB.

The high B^b trumpet, which is also used occasionally for high Baroque parts, no longer has a pronounced main formant, its highest intensity lies at around 900 and 2,000 Hz. Between these, we find a dip in the nasal region. All of this gives the instrument an open clear tone, and also, undoubtedly because of its high voice location, it guarantees a certain security in the attack. Furthermore, inasmuch as the intensity between 3,000 and 5,000 Hz falls by about 20 dB below that of the normal B^b trumpet (for the same total loudness), and also a secondary formant appears around 5,700 Hz, a radiant, bright tone picture is the result, without the hardness, which for the normal B^b trumpet is often unavoidable at that high register.

Score example 4 M. P. Moussorgsky, Pictures at an Exhibition, "Samuel Goldenberg and Schmyle" trumpet part measure 9 ff. (in the orchestration by M. Ravel)

3.1.2.6   The Clarino

The Clarino is a valveless brass instrument, built as a reconstruction of a Baroque instrument, which in its outward appearance is reminiscent of a natural horn. Its basic key of D is located a sixth below the normal B^b trumpet, so that the actual playing region is situated in the region of high order natural tones. The clarino has two overblowing holes, which are intended to increase intonation certainty for high notes.

In the low register the tone is characterized by a formant between 1,000 and 1,200 Hz. This is in the region of the vowel color "a(ah)," and also somewhat below the B^b trumpet (Mühle, 1965). From the fifth octave upward the spectral structure is dominated by the fundamental.

When comparing tones of the same total level, the overtone content of the Clarino (which also has a secondary formant at around 3,500 Hz) is stronger than in the Trumpet because of its greater instrument length (Müller, 1971). However, the individual dynamic steps of the clarino, looking at the intensity, are significantly lower than for the B^b trumpet. The difference for those tones played with both overblowing holes closed amounts to approximately 10 dB; upon opening the overblowing holes the sound level again drops by about 5 dB. In consequence, the absolute *forte* of the trumpet is richer in overtones, when compared to the Clarino, after all. The overall result is a bright, yet soft tone color for the Clarino, which is not as brilliant as that for the trumpet. The initial transient of the Clarino is somewhat longer than for the trumpet, which further supports the soft tonal characteristic.

3.1.3   The Trombone

3.1.3.1   Sound Spectra

The tonal range of the tenor trombone extends to E_1 on the low end, so that the spectral range of this instrument begins at 41 Hz when pedal tones are included. Similarly to the trumpet, for the *mezzoforte* tones of the trombone, the fundamental in the spectrum dominates only in a few cases, rather, the spectra drop below the maximum at a rate of 10 dB/octave for the pedaltones, and 5 dB/octave for the others. The frequency region of highest partial intensity is found in the area of 520 Hz, as shown schematically in Fig. 3.4. This formant location corresponds to the vowel "o(oh)," however, it is more pitch dependent than in the trumpet. While the maximum for low registers is found at 480 Hz, suggesting a clear "o(oh)" sound, in the upper registers the main formant is shifted to 600 Hz, i.e., to a transitional tone color between "o(oh)" and "å(aw)." The sonorous fullness of tone in the low register is transformed into an open, forceful timbre at the high end. However, it needs to be pointed out that the formants of the trombone (with a logarithmic decrement of 2.1) are not as sharply defined as those of the bassoon for example, which for that reason comes significantly closer to the character of a sung vowel.

Above the main formant, the overtone intensity decreases only relatively slowly. At the same time several secondary formants develop. The first one of these contributes particularly to the strength of the striking timbre by emphasizing the components in the "a(ah)" region. Additional weak maxima in the nasal region, as well as in the brightening regions of the vowels "e(eh)" and "i(ee)" complete the tone picture. Furthermore, individual timbre differences can be caused by the instrument itself (bore and bell mouth width), by the player, and by the mouth piece, and this apparently in the order mentioned with decreasing importance (Pratt and Bowsher, 1978).

3.1.3.2 Dynamics

At *ff* the trombone develops a sound extraordinarily rich in overtones, where the 3,000 Hz components are only 5–10 dB below the strongest partials. For an extremely strong attack, this can certainly fall into the region of the second formant. For high frequencies the sound power spectrum drops at a rate of only 3–6 dB/octave. Since the partials in the spectrum are very closely spaced by reason of the instrument's low pitch, a noise-like impression can be caused by the high overtone density at high frequencies. This is perceived as a metallic tone. At low loudness levels only a limited number of partials appear. When compared with the low registers of the horn, this can still be considered as an ample spectrum. Thus the sound power level at around 3,000 Hz lies at about 50 dB below the main formant and furthermore drops by 20–30 dB/octave. The *pp*-tone of the trombone is therefore not very soft, nor is it possible to reduce the overall loudness as much as for a horn.

An excessive weakening of the overtones causes the trombone sound to become dull and lacking in outline. In the low loudness range, therefore, those instruments will sound best, which by reason of a narrow bore are relatively rich in overtones. However, these instruments will then sound hard in *ff* and will not develop a desirable carrying ability. A strong and sonorous *f*-sound is better obtained with larger bore instruments, which, on the other hand, sound less convincing at *p* (Wogram, 1979). This means that the tonal esthetics require less pronounced spectral differences between *f* and *p* than are usually generated by a trombone.

For rapid note sequences one can count on a sound power level of 89 dB at *pp* and 105 dB at *ff*. In the low register individual notes can be reduced down to 73 dB, *ff* notes in the upper register can reach 113 dB. The dynamic range accessible to the performer includes 30–35 dB. The median sound power level for a *forte* lies at 101 dB. The 3,000 Hz components vary by 2.1–2.9 dB for a 1 dB shift of the strongest partials (Meyer, 1990).

3.1.3.3 Time Structure

In contrast to a trumpet, the onset of the tone in a trombone is primarily characterized by the short time span required to reach the final amplitude, rather than by the

sharpness of the preliminary pulse, even for tongued notes. In the low registers the starting transient for tongued notes is about 40 ms, in the upper registers this is reduced to as little as 20 ms (Luce and Clark, 1965; Melka, 1970). Interestingly, trombones reach the steady state in the region of the fourth octave somewhat faster than trumpets. Naturally, for a soft attack the duration of the starting transient is increased; however, with times around 70 ms, the trombone reaches the steady state significantly faster than the trumpet or the horn. For these two instruments the performer obviously has more flexibility to shape the attack than for the trombone.

Figure 3.5 represents two typical pictures for the initial transient behavior of the trombone. In these pictures the amplitude development in the upper representation is to be considered as the normal case, thus as the esthetically more satisfying attack, while the lower picture represents a poorly attacked tone with an excessively strong preliminary impulse. It is clearly noticeable that the amplitude again decreases significantly before the tone finally is developed.

While these pre-cursor tones are frequently pointed out in the literature as characteristic of all brass instruments, it must be noted in contrast, that at least for the lower instruments good players always strive to achieve good tone development. An excessively strong pre-cursor gives a hardness to the trombone sound
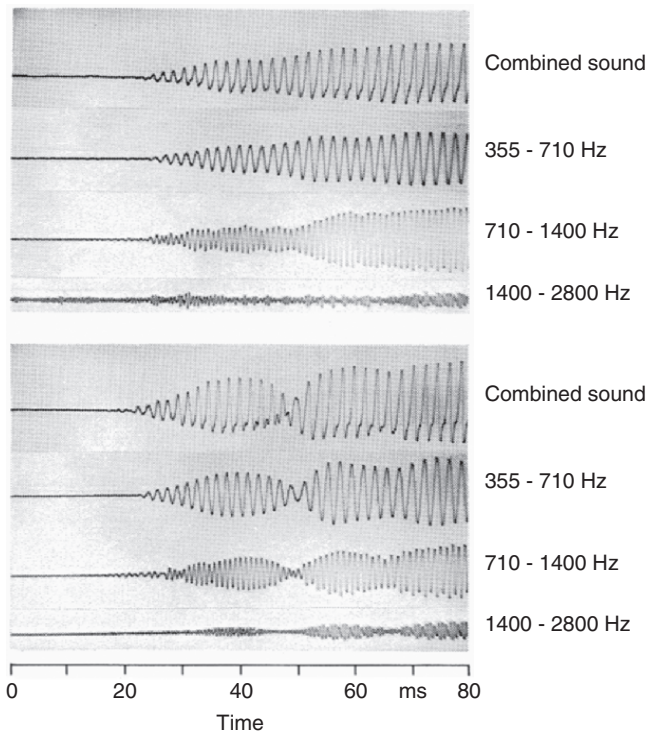

Fig. 3.5 Octave filter oscillogram for initial transient processes in a trombone (played note: B_4^b) *top*: good attack; *bottom*: poor attack

which is mostly disturbing. This can even assume the character of crackling. Nevertheless it is interesting that residuals of a pre-cursor can also be found in a smooth attack. In the frequency region between 710 and 1,400 Hz the vibration amplitude develops in two steps. The first phase displays about the same duration as the pre-cursor in the lower picture. This shows an instrument-dependent tendency for such a pre-cursor effect of roughly 25 ms duration. This corresponds to a round-trip of the sound within the instrument.

In the trombone, in contrast to the other brass instruments, the *vibrato* is not only executed through the lips, but also through movement of the slide. The tonal effect of these two performance techniques is juxtaposed in Fig. 3.6 for the note F_3. The strong time variation of the upper frequency limit is particularly noticeable. Evidently the tone color modulation dominates the frequency modulation.

Typical for a pronounced lip *vibrato* is a *vibrato* width of about ± 10 cents; it causes variations of about 3 dB in the lower partials, this increases to 9 dB at

Fig. 3.6 Time variations of a sound spectrum for a trombone *vibrato* (note F_3) for different playing techniques
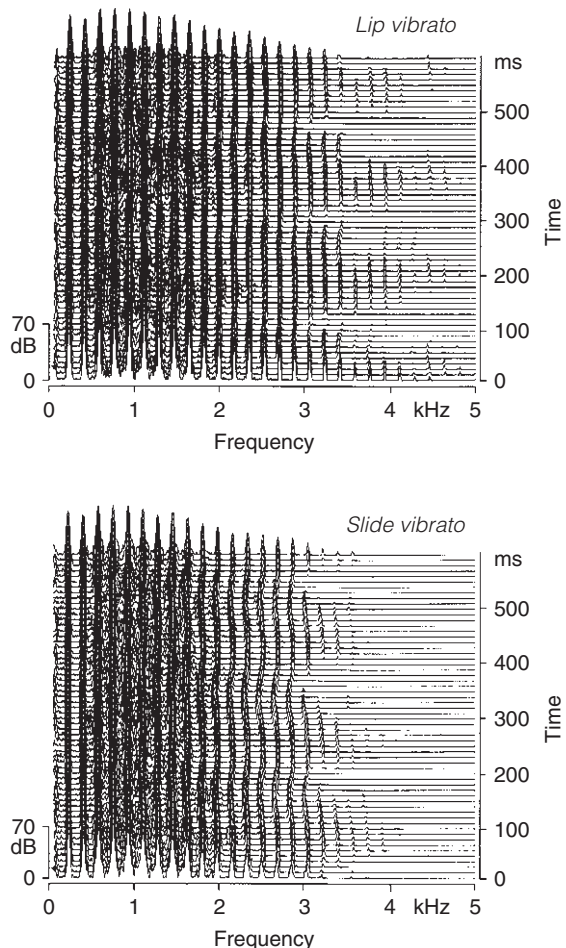
1,000 Hz and to roughly 15 dB at 3,000 Hz. All partials experience this amplitude deviation in phase. For a slide *vibrato* a width of ± 20 dB is not unusual, however, the associated amplitude modulation is less, it only reaches 10 dB at 2,500 Hz. Thus for the slide *vibrato*, the tone color modulation is not as strong as for the lip *vibrato*, furthermore the pitch modulation is frequently implemented through a relatively slow *vibrato* (under 5 Hz.).

3.1.3.4   The Bass Trombone

For orchestral parts specifying bass trombone voices, usually instruments with F valves are used. These, in contrast to normal tenor trombones, have slightly larger bore and sometimes also a somewhat larger bell. This causes the tone to change in the direction of a darker tone color. As example for a very marked bass trombone sound, Fig. 3.4 includes the formant locations for a slide trombone in F (with two valves), which was especially developed for the low passages in Verdi operas, and is therefore designated as "Cimbasso."

    As a comparison with the tenor trombone shows, all maxima are shifted slightly toward lower frequencies, however, the characteristic dense formant sequence of the trombone is maintained. The fundamental formant in the region of the vowel color "u(oo)" provides the substance to the tone so necessary for a bass voice, the actual strength, however, comes to the instrument from the double å(aw)-a(ah) formant which is found in similar form in the Vienna F-horn. The repositioning of the first maximum to about 370 Hz in contrast to 520 Hz for the tenor trombone signifies a lower tuning by nearly a fourth, so that the pitch location and tone color roughly correspond.

    The trombone character of the Cimbasso-sound is evidenced also in the greater frequency range of the spectrum, which was already expressed by the long series of formants. Already for the lowest notes the partials reach up to 3,000 Hz, then, from the second octave on, an overtone series is formed, which also in *mezzoforte* goes beyond 4,000 Hz. The fact that the amplitudes are less than for the tenor trombone, corresponds to the low pitch of the instrument. As the averaged envelopes for the upper region of the second octave show in Fig. 3.7, the difference in partial intensity for equal loudness of tenor trombone and Cimbasso between 1,200 and 3,000 Hz is approximately 10 dB, and above this frequency rises to about 15 dB. The tenor trombone thus has a more brilliant effect.

3.1.4   The Tuba

3.1.4.1   Sound Spectra

The bass tuba and the contrabass tuba are the lowest instruments of the orchestra. The lower limit of their range are at B_0^b (29 Hz) or possibly at A_0 (27.5 Hz), however, these low notes present extraordinary demands for the performer. The
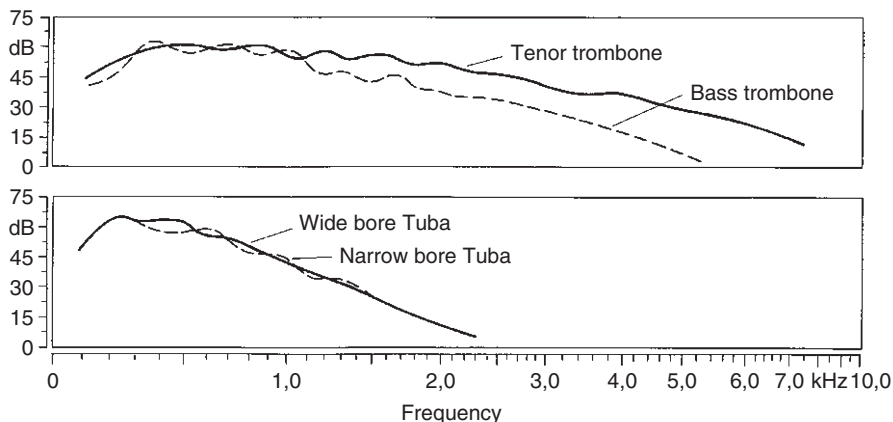
Fig. 3.7 Averaged spectral envelopes of several brass instruments for the tonal region of F_2 to B_2^b

bass tuba in F is naturally located a fifth above the contra bass in B^b, however, it can practically be played equally low, provided its bore is sufficiently wide (Kunitz, 1959). Thus the essential difference between a bass tuba and a contrabass tuba rests in the sound, ignoring the seldom needed upper region of the playing range.

The spectra of the tuba differ from those of the trombone in the low register mostly in the significant decrease of overtone content. For low notes at *mezzoforte,* the upper frequency limit for harmonic contributions lies between 1,000 and 1,500 Hz, depending on the structure of the instrument. In the middle and upper registers the partial series is broadened to frequencies between 1,500 and 2,000 Hz. This steep amplitude drop for higher tone contributions is clearly expressed in the spectral envelope of Fig. 3.7, showing a characteristic totally different from trombones. At the same time they also show that the lower partials are significantly weaker than the strongest components. Below the 1st formant, the power spectrum of a tuba drops by 10–15 dB/octave depending on the bore.

The main formant, which is thus generated between 210 and 250 Hz depends somewhat on the design, for the wide instrument it is lower than for the narrow one. As Fig. 3.4 shows, the tone color for both cases is somewhat darker than a normal "u(oo)". For the wide bore instrument a secondary formant is located on the boundary between "u(oo)" and "o(oh)," while for a narrow bore the corresponding secondary maximum is related to an open "o(oh)." Inasmuch as, particularly in the upper registers, there are no additional secondary formants for the wide bore instrument, the narrow tuba produces a somewhat slimmer timbre, while the wide bore emphasizes the dark, soft and occasionally muffled tonal impression.

3.1.4.2 Dynamics

The *pp* tone of a tuba is determined by a spectral decrease with a slope of 20 dB/ octave above 250 Hz, it is thus very dark and soft. In contrast, above the actual

overtones, disturbing noise components become noticeable at greater loudness levels. As a result the tone should not be forced, lest it become raw. Esthetically determined upper limits should therefore take precedence over technically attainable upper limits for *ff*. With that in mind, Richard Strauss (1905), in his text on instrumentation, specifies that the tuba not play beyond *mf*, and it is therefore very understandable that Berlioz, in the last movement of his Symphonie fantastique, has two bass tubas play the "Dies Irae" in unison.

The sound power level of the tuba at *forte* lies at 104 dB; rapid note sequences can vary between 93 and 108 dB. Low register individual notes can be considered as relatively subdued at a sound power level of 77 dB, particularly when considering the reduced sensitivity of the ear at low frequencies. Generally, a playable dynamic range of 25–30 dB is obtained, where the 3,000 Hz components change less than for other brass instruments, i.e., only by a factor of 1.5–2 of the amplitude change for the strongest tone contributions. The technically possible *ff* of 112 dB in the high register is in practice of no significance (Meyer, 1990).

3.1.4.3 Time Structure

In spite of the low pitch, a very fast initial transient is achieved in the tuba. The duration of the starting transient for tongued notes in the region of C_2 is about 40 ms. For the upper registers this is reduced to as short as 25 ms. For the lowest registers, however, this is increased to over 60 ms (Luce and Clark, 1959; Melka, 1970). If a *staccato,* even for a good player, does not sound as crisp as for other brass instruments in spite of the short initial transient, then the cause is found primarily in the fact that the tone is not rich in overtones. Furthermore, the preliminary impulse is not as steep in its rising slope as for other brass instruments, so that it does not characterize the attack as precisely.

Surprisingly, softly attacked tuba tones reach their steady state more rapidly than is the case for horns and trombones. While the initial transient for C_1 can last for over 130 ms, corresponding values for higher registers lie around 60 ms, and again the limited frequency range of the spectrum supports the soft tone onset. This also explains why the overtone-rich trombone requires a slower initial transient process. The horn, on the other hand, already occupies a special position with the particularly wide modulation region for its tone development.

3.2   Woodwind Instruments

3.2.1   The Flute

3.2.1.1 Sound Spectra

The tone of transverse flutes is characterized by the very uniform overtone structure of the spectra. With few exceptions, particularly for the notes between C_4 to or E_4^b or E_4, the fundamental is the most strongly developed of all partials for the entire

range of the instrument. For no other orchestral instrument is this characteristic as clearly marked. Above the fundamental the intensity of the overtones drops quite linearly with increasing frequency. Secondary formants appear only occasionally, and then very weakly. They are thus not typical for this instrument group, but rather characterize tonal idiosyncrasies of individual instruments.

The intensity relationship between the lower partials, and thus the tone color, can be varied by the performer within relatively wide limits. The performance technical parameters available are the blowing pressure (i.e., the air stream velocity through the lips), the degree of coverage of the mouth hole (and thus the distance between lip opening and blowing edge) and the blowing direction. The size of the lip opening affects only the dynamics and not the tone color. Raising the blowing pressure leads to a better coincidence of the first overtones with their associated resonances, increasing their strength relative to the fundamental. The tone thus brightens in its color, which supports the impression of the changed timbre caused by the dynamic change resulting from the increased pressure. Reduced coverage (increasing the distance s in Fig. 3.8), in addition to influencing the intonation, leads to a softening of the air stream at its edge. This causes a weakening of the overtones for unchanged fundamental strength; the shorter the distance the brighter will be the tone color. Furthermore, by varying the coverage, the intonation of the higher resonances can be adjusted, within limits, to minimize noise contributions. Finally, the blowing direction determines the strength relation between even and odd numbered partials. As Fig. 3.8 shows, an air stream directed symmetrically toward the blowing edge ($y = 0$) enhances the fifth, while for a slightly more outward or inward directed air stream the octave or double octave is strengthened and the fifth is diminished. The flute sound is perceived as having particular tonal beauty when the fundamental and octave are approximately equal in strength, and the twelfth is weaker by about 10–15 dB (Bork and Meyer, 1988).
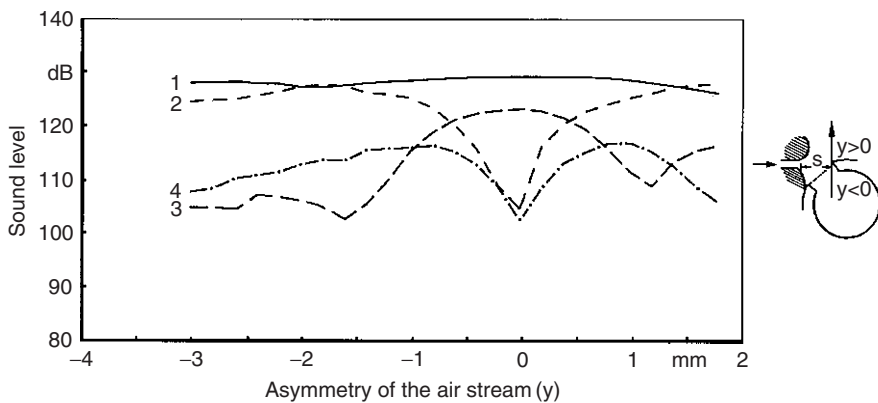

Fig. 3.8 Strength of the first four partials in the sound spectrum of a flute (played note: C_4) in their dependence on blowing direction

As already suggested in Fig. 2.6, the noise in the tone of a flute consists not only of flow noise with intensity independent of frequency, but it also contains components which influence the tone. They appear in overblown notes and result from continuous statistical excitation of "unused" resonances. In the region of E_5 to D_6, therefore, such flow noise peaks appear at one half of the fundamental frequency and odd multiples thereof; also from E_6^b on up, at 1/3 and 2/3 of the frequency of the fundamental and at the corresponding multiples not divisible by three; in particular, for cross fingering, the flow noise peaks can also be inharmonic. The height of these peaks depends largely on the quality of the instrument and the playing technique. Contrary to a widespread opinion, for metal flutes the material appears to play a subordinate role in relation to tonal impressions, especially for the listener (Coltman, 1971). Special "preferences" for gold or even platinum flutes rest predominantly on a psychological effect, which of course, in individual cases may certainly influence the performance technique of the flautist.

3.2.1.2 Dynamics

As already mentioned, the player can influence the dynamics within relatively narrow limits (about 6 dB) without changing tone color, by variation of the lip opening. The predominant dynamics contributions, however, are accompanied or supported by changes in tone color. In contrast to most other instruments this is above all relevant to the lower partials, and less so in the upper frequency region. Thus at *ff* the 3,000 Hz components lie about 20–30 dB below the fundamental, and for *pp* only by 30–40 dB. Above that, all spectra drop with a slope of 15–22 dB/ octave. The dynamic range of the flute is relatively small. For rapidly performed scales through two octaves it is only 12 dB with a sound power level of 82 dB at *pp* and 94 dB at *ff*. This is also related to the fact that for a flute the playable dynamic range depends especially strongly on pitch. The low register with a *pp* around 67 dB and a *ff* in the neighborhood of 86 dB is particularly weak; in the highest octave the sound power level can reach 100 dB for *ff*, while for *pp* it can hardly be lowered below 83 dB. When one further takes into account the low sensitivity of the ear at low frequencies, then these values suggest that a low frequency *ff* is perceived as equally loud as a *pp* of the highest notes. Consequently, for a flute, the tonal time structure receives particular significance by its influence on tonal emphasis.

As a whole one can count on a practically realizable dynamic range of from 15 to 20 dB in a flute. The average *forte* power level lies near 91 dB. A change of the fundamental by 1 dB is connected – depending on performance technique – to a change of the 3,000 Hz components of up to 1.5 dB in the lower and middle register and 2 dB in the upper register (Meyer, 1990).

3.2.1.3 Time Structure

Among wind instruments, the flute requires the longest time for tonal development. In addition, the attack contains particular characteristics which appear in response

to initial blowing techniques. Dominant among these are so-called preliminary tones, which are formed by higher resonances (Rakowski, 1966).

For the lowest notes, these preliminary tones have a duration of about 50 ms. They are located in a frequency region around 2,000 Hz, that is approximately three octaves above the fundamental. Their intensity is about 10 dB greater than the subsequent stationary state. For the lower partials it is typical for the tone picture to have the octave partial exhibit a relatively fast initial transient and the actual fundamental to develop substantially more slowly thereafter.

In the mid-range the initial transient is as fast for the fundamental as for the next higher tone contributions. In addition to the preliminary tone, which raises its frequency to 4,000–6,000 Hz in correspondence to the three octave rise above the fundamental, noise-like components also appear below the fundamental, which disappear after about 100 ms. The duration of the high preliminary notes is of the order of 50 ms. Finally, the initial transients for high overblown notes consist initially of sub-resonances before the energy passes to the actual partials after about 50 ms.

The strength of the preliminary tones naturally depends on the sharpness of the attack, or on the articulation syllables used; however, it should by no means be considered as a negative characteristic, to be suppressed by performance technique. It enhances the sharpness perception of the attack, which is especially important in view of the relatively long initial transients for flute staccatos (low register about 100 ms, mid-register around 30 ms.). This was already shown very clearly in Fig. 2.9, the juxtaposition of a sharply articulated and a soft tone attack. Finally, preliminary tones, as well as noise components generated by the attack, participate substantially in the tone effect known as the "flutter tongue," as used, for example, by R. Strauss in characterizing the wind in tone painting fashion in Variation VII of Don Quixote (see score example 5). Here the individual tongue beats of the "flutter tongue" follow each other with a frequency of 25 Hz, and for higher frequencies lead to an amplitude modulation of 15 dB.

While the end of a tone for brass instruments and reed instruments is characterized by a very short decay time, because of the sudden cessation of lip or reed vibration, the decay process for a flute can be influenced somewhat by the performer. This is unique among wind instruments. For a "normal" flute tone ending in the mid register, the decay time (above 60 dB) is at about 125 ms for the fundamental and 80–100 ms for the next three overtones. A soft termination of the tone prolongs the decay time to about 200 ms for the fundamental and the octave, and to 120 ms for the following overtones. Naturally, even these values are short in comparison with string instruments.

Of particular importance for the flute sound is the effect of the *vibrato* on the spectrum. Two characteristic examples are shown in Fig. 3.9. The width of the *vibrato* for a flute is relatively small. Frequency fluctuations of from ± 10 to ± 15 cents can already be considered as a strong *vibrato*. For this amount there are practically no amplitude variations in the lower partials; the effect therefore is one of a stable tone. Only the higher overtones move to the resonance wings by reason of the *vibrato* and therefore experience strong amplitude modulation. Above 3,000 Hz this level

Score example 5 R. Strauss, Don Quixote, Variation VII, flute passages
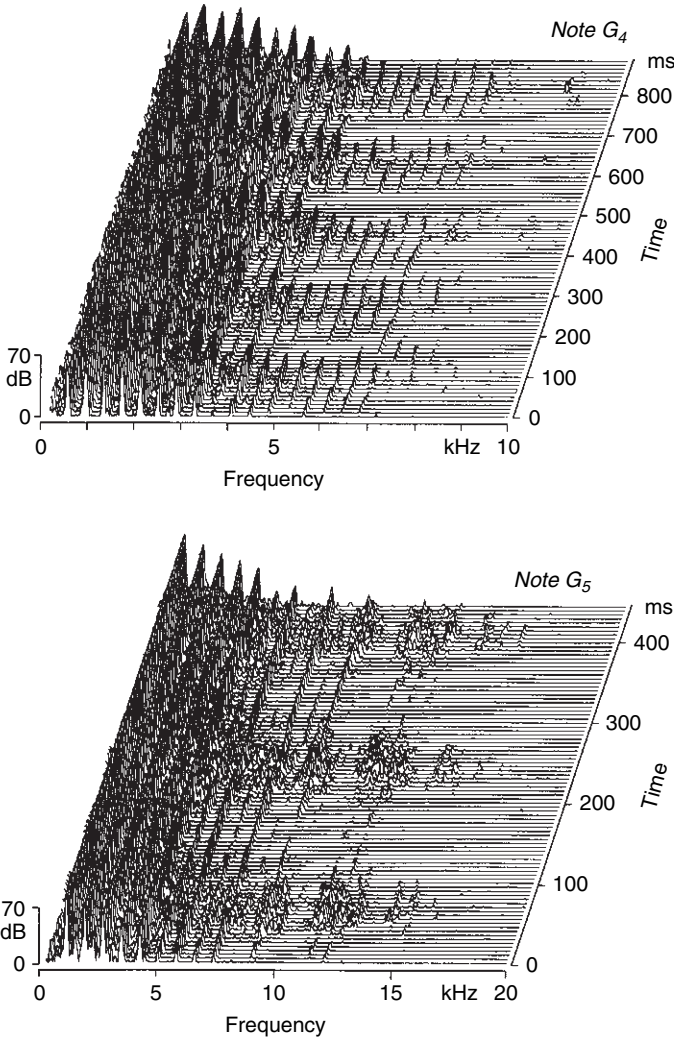
Fig. 3.9 Time variations in the tonal spectrum of a flute vibrato for different pitch regions

fluctuation reaches 15 dB and effects a pulsing tone color variation. For example, this effect is clearly pronounced for Note G_4 and also recognizable for Note G_5. Added to this is a very pronounced modulation of the blowing noise which reaches well above 10,000 Hz with a level fluctuation of 15 dB. The tone color modulation, as well as the noise modulation, lend the fluctuating flute tone a particular prominence within an ensemble, even when the blowing noise is not as strongly pronounced as in the pictured example (Meyer, 1991).

3.2.1.4 The Piccolo

As is the case for the flute, the piccolo in its spectrum also shows a steady amplitude drop beginning at the fundamental, which is the strongest partial, to the higher tone contributions. In this context, a secondary formant can appear in individual cases in the region around 3,000 Hz, which, corresponding to its location in the vowel "i(ee)" region, supports the very bright tone of this instrument. In the middle and upper regions of the tonal range, harmonic components are formed up to about 10,000 Hz which nevertheless do not give metallic sharpness to the tone since the frequency separation of the partials for the piccolo is very large, due to the high pitch. The timbre is better described as bright and penetrating, which is particularly connected to the fact that the strongest tone contributions of this instrument fall precisely into the frequency range of highest sensitivity for the ear.

In addition, it should be noted that the dynamic performance range of the piccolo is particularly narrow. On the average it barely covers 15 dB, and it is just that wide over the entire tonal range (Burghauser and Spelda, 1971). At *pp* the piccolo is therefore very loud in comparison to other instruments: the lower dynamic limit rises from a sound power level of 78 dB for the low notes, to a value of 88 dB in the highest register. Correspondingly an *ff* produces values between 93 and 103 dB, so that in high passages the piccolo can certainly be made to stand out in a full instrumentation orchestra sound: Particularly typical examples of this are found in the symphonies of D. Shostakovich.

3.2.2 The Oboe

3.2.2.1 Sound Spectra

As determined by the different process of generating vibrations and the conical nature of the bore, the oboe has a totally different tonal character than the flute. Acoustically this is borne out in the spectrum which is rather rich in overtones. The frequency distribution of this spectrum is determined by a series of formants. Even though the spectral region reaches down to about 233 Hz (the fundamental for B^b), the strongest tonal contributions do not appear until about 1,100 Hz. This principal formant, as shown in Fig. 3.10, leads to a basic tone color similar to the vowel "a(ah)." In addition, for the lowest notes a certain sonority is obtained by a subformant between 550 and 600 Hz which can shift the tone color in the direction of a bright "o(oh)."

The fundamental of the oboe sound is relatively weak especially in the low registers, since the sound power level below the formant maximum drops by 4–6 dB/octave. The fundamental thus lies up to 15 dB below the main formant. Its vowel character is therefore enhanced. Two secondary maxima at around 2,700 Hz, i.e., in the boundary region between the colors "e(eh)" and "i(ee)," and around 4,500 Hz lend a pronounced brightness to the oboe, which is also enhanced by further partials up to 9,000 Hz (at *mf*). Finally, in the low register, the frequency location of the sub- and main formant results in the effect that the even partials
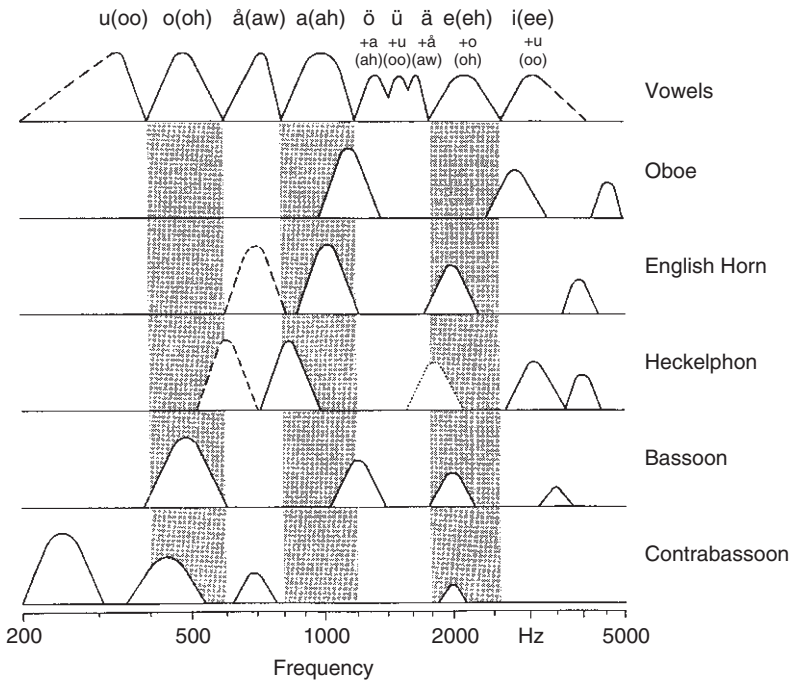
Fig. 3.10 Schematic representation of formant locations for double reed instruments

dominate in intensity over the odd partials. This differentiation (of the order of 8 dB) leads to a particularly open sound.

At the boundary for overblown tones, the formant region becomes discontinuously broader, above D_5 the formant thus loses in character (Smith and Mercer, 1974). Up to about B_5^b the fundamental and octave are nearly equally strong, so that the maximum no longer exhibits a clear peak, or the octave partial may even dominate somewhat. From B_5 on up the fundamental dominates; for the higher registers, therefore, the vowel-like tone picture of the oboe increasingly gives way to a rather hard and less expressive timbre. How critical these high notes of the oboe are in view of tonal considerations is noted, for example, when in the slow movement of the G major violin concerto of W. A. Mozart, the voice carried up to the D_6 by the flute is taken over by the oboe, as happens occasionally, because in the outer movements only oboes are present.

The typical tonal difference of Vienna oboes in contrast to generally played instruments of the French style, rests in the fact that because of the somewhat narrower cone, the main formant lies by 50–100 Hz lower, and the higher overtones are more pronounced. Thus the tonal character becomes somewhat more pointed and less nasal. For Baroque oboes the main formant, depending on instrument construction, lies more or less far below 1,000 Hz, so that the tone region determined by this formant is stretched by several half steps in comparison to today's instruments (Benade, 1976).

3.2.2.2 Dynamics

When considering dynamics, varying the tone color does play an important role for the oboe, however, the nature of reed vibrations imposes limits. While the *ff* for brass instruments is determined by strong tone components even above 10,000 Hz, the center for change in the power level spectrum of the oboe lies in the region around 3,000 Hz: for the oboe *ff*, these components are located only 10 dB below the strongest tone contributions; above 3,000 Hz the envelope drops relatively steeply with 23 dB/octave. For *pp* the linear amplitude drop already begins at 1,500 Hz and has a slope of 20 dB/octave or less, so that the 3,000 Hz component lies up to 30 dB below the strongest tone contributions. As a result, for *pp*, tones are produced with only very few strong partials: In the low register the spectrum is reduced to four to five harmonics, where the amplitude maximum is shifted to the second partial (in contrast to the 4th or 5th in *mf*). In the upper registers, actually only two partials appear, and the fundamental dominates in intensity by far. For *pp* the main formant is thus shifted in the direction of a somewhat darker basic color, in addition, the tone becomes more tender due to the lack of higher frequency contributions.

In comparison to the influence of the tone color on the dynamic range, the difference between the power level at *ff* and *pp* is relatively narrow. For rapidly performed scales the level at *ff* rises to 95 dB in contrast to 83 dB at *pp*. For isolated notes the power range expands to values from 70 dB for *pp* in the mid range to 103 dB for *ff* of very high notes. One can thus count on a practically usable dynamic range of up to 30 dB, with only 20 dB in the extreme registers. For the characteristic value of the oboe *forte* sound, a sound pressure level of 93 dB can be given. The "dynamic tone color factor" which is responsible for the relative shift of the 3,000 Hz components moves between 1.7 and 1.9, and is thus a little smaller than for flutes and clarinets.

3.2.2.3 Time Structure

The attack for double reed wind instruments is distinguished by particular precision and clarity. The reason for this lies in the facts that on the one hand initial transients are extraordinarily short, and on the other, that there are no noise-like or inharmonic contributions. In fact, in the initial transient, the individual partials experience a nearly exponential amplitude development so that the auditory impression due to the smooth envelope is clear and pure. Oboes are therefore especially suited for a short, and yet tonally precise *staccato*, which can lead to difficulties in performance alongside other instruments which are not in a position to give their *staccato* passages the same pearlescent clarity.

The initial transient for tongued notes, even for the lowest notes of the oboe, is shorter than 40 ms and is lowered to less than 20 ms with increasing pitch, where the high frequency contributions will already have reached their final strength in about 10 ms. Yet, with "*cantabile*" playing, tones can also be developed softly.

In this, the initial transient can be stretched to 100 ms and can last in the upper registers for about 40 ms (Melka, 1970). Thus the initial transients for the oboe with soft attack can reach similar values as the violin for sharp attack. This comparison makes the influence of the initial transient on the tone picture of the instrument especially clear. Its complement is found in the short decay time of the oboe which lies in the order of 0.1 ms.

3.2.2.4 English Horn and Heckelphone

As lower instruments in the oboe group, the English Horn, and for particular compositions also the Heckelphone, are used in the orchestra. Corresponding to its tuning in F, the English Horn is located a fifth below the oboe, while the Heckelphone lies an octave lower than the oboe. Both instruments possess a pear-shaped bell, which is especially important for the characteristics of the lowest notes.

From a tonal standpoint, the English Horn and the Heckelphone represent a transposition of the oboe timbre into a lower range, as can be seen from the survey of formant regions in Fig. 3.10. In the oboe the main formant can be described as a bright "a(ah)," in the English Horn the vowel color of a dark "a(ah)" dominates and in the heckelphone a transition color between "å(aw)" and "a(ah)."

This formant shift relative to the oboe, however, is not as strong as the shift of the tonal range. In contrast to the pitch shift by a fifth, the shift in color only involves about a whole step, while the heckelphone main formant lies a third below that of the oboe.

In the lower registers of the tonal range an additional formant comes in, which in the English horn corresponds to a somewhat brighter "o(oh)" and in the Heckelphone to a darker "o(oh)." This generated the basic color for the "rather wailing" tone of the English horn and the relatively dark timbre of the heckelphone.

As the location of the higher formants in Fig. 3.10 proves, the English horn and the heckelphone form a decided expansion of the oboe group in the direction of lower pitch. The continuity of this tonal line is illustrated particularly impressively by the score excerpt from "Salome." On the other hand a comparison of formant locations also permits recognition of the difference between the heckelphone and the bassoon from a tonal standpoint. In addition, it is noted that these two lower oboe instruments do not require more initial transient than the oboe itself (Meyer, 1966a). However, the dynamic range in comparison to the oboe is narrower; for the English horn the limits for the sound power level lies near 79 dB at *pp*, and 94 dB at *ff*.

Score example 6 R. Strauss, Salome, Part excerpt: Oboe group

3.2.3 The Clarinet

3.2.3.1 Sound Spectra

The clarinet presents a typical example for sounds, for which the odd partials outweigh the even ones (Backus, 1961 and 1963; Meyer, 1966c; Strong and Clark, 1967). However, they do not preserve these characteristics over the entire pitch range, they change their spectral composition in the region of the upper register in favor of the even tone contributions. Fundamentally the playing region of the B^b clarinet can be divided into three segments, which show different characteristics in relation to their spectral construction and thus their timbre. The boundaries between them are somewhat fluid and depend to some extent on construction details, reed strength and performance technique. Fig. 3.11 contains a juxtaposition of three typical sound spectra for the individual regions, which, however, show the dominance of the fundamental over all other partials, as an essential common characteristic.

In the low octave of the range from D_3 (147 Hz) to D_4 the odd partials are significantly stronger than the even ones. The difference can be followed for the lowest notes up to about the 15th partial; the second and fourth partials are particularly weak in their amplitude, the difference between them and the neighboring odd partials is generally greater than 25 dB, it can even take on values up to 40 dB. As a result the tone becomes dark and hollow; in this register the clarinet is therefore especially suited for "achieving of dark and sinister tonal effects," as for example at the beginning of the first movement of the 5th symphony by P. I. Tchaikovsky (Kunitz, 1957).
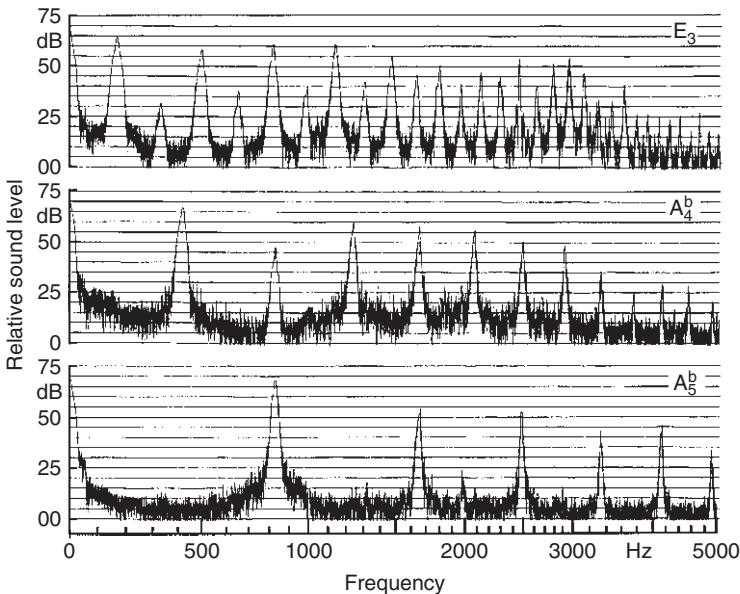

Fig. 3.11 Sound spectra of a clarinet

In the region from approximately E_4^b to G_5 the 1st and 3rd partials are decidedly more strongly pronounced than the octave partial, from the 4th harmonic on upwards the odd and even contributions are, however, equally developed. For a good player, the overblow boundary, which is located near A_4, cannot be recognized from its spectral construction. Nevertheless, the so called short notes up to $G_4^\#$ can occasionally sound somewhat dull, when the amplitudes above 3,000 Hz are too weak. In this pitch range the indifferent location of the octave partial also plays a role: with increasing pitch its difference from the 1st and 3rd partial diminishes, so that cases can arise where it is too strong for a decidedly hollow timbre, and too weak for a forceful sound.

Above $G_5^\#$ the fundamental dominates in a particularly strong measure. It is associated with a steadily decreasing overtone series, which gives a "full round substance" to the high register (Kunitz, 1957). Luster and brilliance are additionally achieved by a formant between 3,000 and 4,000 Hz, i.e., in the region of the vowel color "i(ee)." The large intensity difference between the harmonic partials and the noise background is responsible for the clarity of the sound. It gives the tone an especially pure sound. This becomes particularly clear in comparison to the flute, which in this pitch range exhibits a similar overtone structure, but produces considerably higher noise contributions, and furthermore presents stronger fluctuations in the temporal micro-structure of the sound level, so that the tone of the clarinet gives the effect of being more steady, and firmer and thus also stronger.

3.2.3.2 Dynamics

Of all wind instruments, the clarinet can produce the softest *pp*. At that point the power level drops to about 65 dB and in the region of the fifth octave can even be lowered to 57 dB. The result of this is a sound pressure level, which in a larger hall approaches a value with an order of magnitude around the ambient background noise. For fast scales the *pp* power level does rise to 77 dB. In the highest registers, like above D_6, such a pronounced *pp* can no longer be blown.

For *ff* the power levels for rapid note sequences reach a value of 97 dB. Individual notes, especially in the fifth octave, can rise to 106 dB. This results in a dynamic range, which in its breadth is hardly found in any other instrument. In the lower registers it is measured at around 30 dB, in the mid-registers at a scant 50 dB, and in the upper registers at about 25 dB. The characteristic *forte* for the clarinet lies near 93 dB.

The extraordinarily wide-spanned dynamic expression possibilities of the clarinet are further enhanced by the strong tone-color variations existing between dynamic steps. (Meyer, 1966c). For *ff* the intensity maximum of the low notes is shifted to frequencies above 1,000 Hz, i.e., into the region of the bright a(ah)-formant. For lower frequencies the spectrum drops with 3 dB/octave, so that in this exceptional case the fundamental does not dominate. While in the low register the power level spectrum drops with a slope of 12 dB/octave above 2,500 Hz, in the mid- register the very strong partials persist up to about 5,000 Hz, followed by a

spectral drop of 23 dB/octave. This richness in overtones effects a pronounced brilliance of the *ff*; however, this can also present a certain hardness in the low registers, which is associated with the fact that the amplitude difference between even and odd partials becomes larger with increasing performance strength.

In contrast, for *p* the clarinet tone becomes softer in character through the intensity equalization between the two partial groups. While the spectra in the low register, even at *pp,* are still relatively rich in overtones, since above 600 Hz they drop only by 15 dB, in the fourth octave decidedly tender tones can be produced with only three or four partials, i.e., a spectrum of less than 1,500 Hz. The dynamics are therefore predominantly determined by the overtone content. The partials in the frequency region around 3,000 Hz lie from around 40 dB (low register) to 50 dB (high register) below the fundamental for a *pp*, while the level difference at *ff* only amounts to 10–12 dB. At the same time the "dynamic tone color factor" rises from 2 for the low register to above 2.5 for the upper registers.

3.2.3.3 Time Structure

As is the case for the double reed instruments, the attacks with the clarinet can be very clear and precise. Here, staccato notes practically do not contain higher frequency preliminary sounds – as was the case, for example, with the flute, but rather exhibit a quite uniform amplitude growth in all overtone regions. As Fig. 3.12 shows for the lowest note of the clarinet, even the fundamental reaches its final strength within a few vibrational periods. Thus the attack does not give the pointed effect of the oboe. As a whole, the initial transient is completed 15–20 ms after the attack, yet it can be stretched to more than 50 ms for a soft attack. These values are valid for the entire pitch range of the clarinet (Melka, 1970). The decay time, even in the low register, is not longer than 0.2 s and drops to about 0.1 s in the upper register.

In addition to the slower initial transient of the lower partials, the features of the soft attack are above all formed by the fact that higher overtones increasingly delay the initial transient with increasing frequency. Thus, special importance goes to the transfer of the high tonal contributions for differentiation in articulation, even though, as in the flute, they do not dominate the transient process.

Occasionally noise-like components appear in the attack, when closure of the flaps triggers resonances which have not yet been excited by reed vibrations. This phenomenon plays an important role, especially for connected note transitions.

In the clarinet, a *vibrato* can be produced either with the diaphragm or through the lips. In the rare event that a clarinet *vibrato* is used in classical music, the associated frequency variations are only slight. For a diaphragm *vibrato* a width of about ±7 cent would be typical, this is thus at the limit of audibility for frequency modulation. A lip *vibrato* can reach a width of ±15 cents in the upper registers. These frequency fluctuations are connected to amplitude changes of higher partials, particularly at frequencies around 2,000–3,000 Hz, which can amount to up to 20 dB, however these are, at least partially out of phase. Thus their effect in the ear
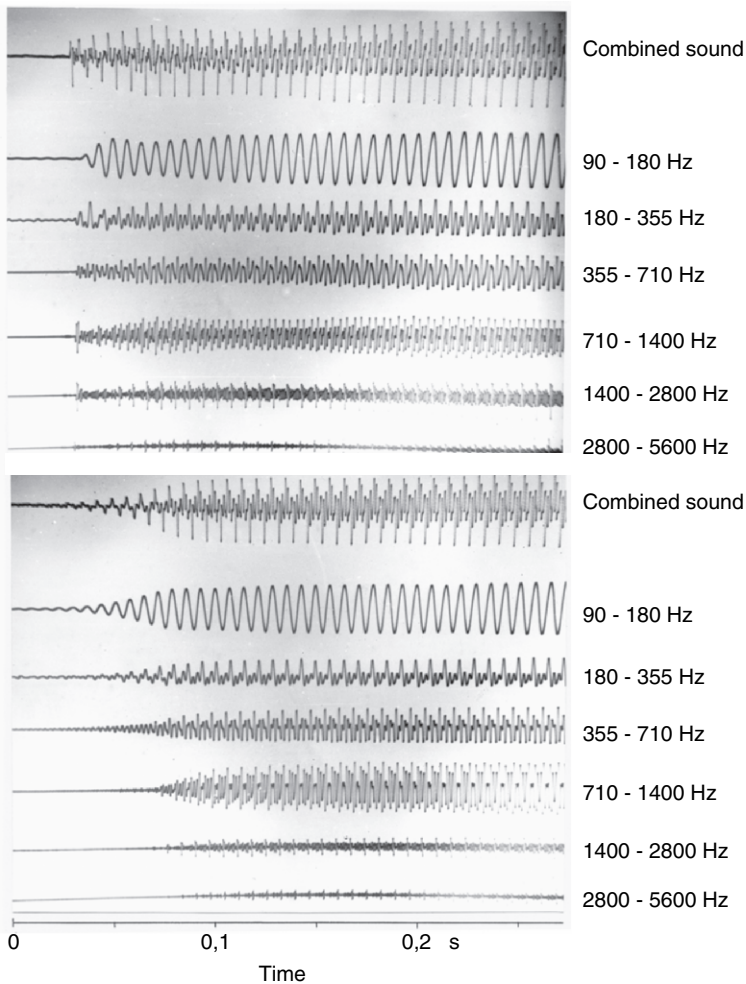
Fig. 3.12 Octave filter oscillogram of initial transient processes in a clarinet (played note: D_3, sharp and soft attack)

is compensated to a lesser extent: within individual frequency bands, the energy of the partials varies only by 5–6 dB for a diaphragm *vibrato* and by 7–8 dB for a lip *vibrato*. A lively tone color effect is thus created in the tonal impression of the clarinet *vibrato*, which overshadows the effect of frequency modulation.

3.2.3.4 Clarinets of Different Pitch

The influence of the instrument's pitch on the tonal character is particularly pronounced for the clarinet (Meyer, 1966c). Above all the difference between an A and a B^b clarinet rests on the fact that – especially in the low register – the intensity of

the A-clarinet in the overtone region around 1,000 Hz, which is in the region of the vowel color "a(ah)," is about 5 dB lower, also the contributions above about 3,000 Hz are weaker. A further nuance arises from the fact that in the A-clarinet the even partials are even more reduced. All these phenomena lead to a very "dark and song-like" timbre, which at times is also specified as "holding back and tender" (Kunitz, 1957). In contrast, the B^b-clarinet gives a more brilliant and forceful impression as a result of its somewhat richer partial spectrum. It is noteworthy, how much even small intensity differences matter in differentiating in the tonal picture, they are nevertheless sufficiently significant for Richard Strauss to employ both instrument types at the same time in Salome, the A-clarinet for melodic passages and the B^b-clarinet for brilliant figures and ornaments.

While for the A- and B^b-clarinets a dominance of the odd harmonics up to about the notes $F_5^{\#}$ or G_5 is observed, for the clarinet in C this boundary is shifted up to about B_4^b; thus it coincides with the over blow boundary. Furthermore, the intensity difference between even and odd partials in the low register of the instrument is less than for the larger clarinets. Thus its tone color is not as hollow and covered. The C-clarinet exhibits significantly higher amplitudes in the region between 1,500 and 4,000 Hz, the difference in relation to the lower instruments is about 10 dB and effects not only a significantly brighter sound, but because of missing formants a "cooler and harder" timbre. These characteristics render the C-clarinet particularly suitable for folklore related tasks such as, for example, the polka in the first act of the "Bartered Bride."

Still more strongly pronounced is the difference between the clarinets in D and E^b, and the normal clarinets. Already from 1,000 Hz upward much higher amplitudes appear in the high clarinets, which in the region around 2,500 Hz make a difference of more than 25 dB, in contrast to the larger clarinets. Inasmuch as the level difference between even and odd partials shrinks to 10–15 dB, and because of the diminished intensity drop above 1,000 Hz, the nasal components are relatively strongly developed, a bright, frequently even shrill tone color results, which on occasion can have hard and pointed effects. The high clarinets are thus predestined, above all, for special tonal effects, as they are demanded in "Feuerzauber" (Walküre), or "Eulenspiegel." The dynamic range of the small clarinets is relatively narrow, with median power levels of 80 dB at *pp* and 96 dB at *ff*, a range of only 16 dB is noted. The possibilities for expression are thus significantly less – especially for *piano* – than for normal clarinets.

The bass clarinet is characterized by properties already considered in the A-clarinet. In particular, the difference between even and odd partials grows to more than 30 dB, where the fundamental is especially emphasized. Through this the timbre becomes particularly hollow and dark, it obtains, as it were, something "mysterious and melancholy." The *pp* can be played extremely discreetly, in the lower register it drops to a power level of 59 dB, where for subjective tone impressions one must still consider that the loudness impression is further reduced by the decreasing aural sensitivity at low frequencies. At *ff* the instrument is also not very strong, a power level of only about 97 dB can be expected.

3.2.4   The Bassoon

3.2.4.1   Sound Spectra

At the low end, the range of the bassoon extends to B_1^b (58 Hz), with an extended bell joint A_1 can be reached, which for example is required for "Tristan und Isolde." The lowest frequency contributions in the spectrum of the lowest notes, however, are developed only relatively weakly. The intensity maximum is not formed until the 8th or 9th overtone. This frequency location corresponds to the region of the main formant, which lies at about 500 Hz, and which particularly for bassoons is very pronounced. (Lehman, 1962; Meyer, 1966c; Strong and Clark, 1967).

   The tone color which is very similar to the vowel "o(oh)" is not only caused by the central location of the amplitude maximum in the region of this vowel, but also rests on the fact that the width of the bassoon formant nearly corresponds to that of the vowel "o(oh)." For all other instruments the formant width is wider than the width of the vowel with corresponding frequency location. The logarithmic decrement has a value of 1.4 (Meyer, 1968) for the prevalent construction of a German bassoon, the corresponding value for a sung "o(oh)" is 1.2 (Tarnóczy, 1943). A fitting example for the dramatic exploitation of this similarity of the bassoon with the human voice is given by the duet of Aida and Amonasro from the Nile scene in the opera Aida (see score example 7), where also during the singing of Aida the presence of her father is given musical expression.

   This formant location around 500 Hz characterizes the tones of the bassoon in the region up to C_4. Below this formant, the power spectrum drops with a slope of approximately 8 dB/octave, so that the fundamentals of the lowest registers are developed correspondingly weakly. In the fourth octave the amplitude maximum is shifted along with the 2nd partial to somewhat higher positions, so that the tone color approaches an "a(ah)." From about B_4 on up, the fundamental dominates, so that the vowel color is no longer as pronounced.

Score example 7   G. Verdi, Aida, 3rd Act, Duet Aida – Amonasro, score excerpt

Additional tonal characteristics of the bassoon include a very high number of overtones, which form a series of secondary formants in several strong groups. As already recognizable from Fig. 3.10, and represented again by Fig. 3.14 in an other context, these secondary formants lie in the regions of 1,150, 2,000, and 3,500 Hz. Of these, the first one falls into the region of the vowel color of a bright "a(ah)" which contributes to a strong timbre. In the higher regions of the tonal range it is shifted more into the region of the nasal components, so that the tone picture no longer gives such an open impression. A typical example for this somewhat nasal color of the bassoon in the upper registers is given by the Beckmesser-motif in the "Meistersinger" (see score example 3, p. 48 Sect. 3.1.1.4) The two higher frequency formants account for a tonal brightening effect corresponding to their position near the vowels "e(eh)" and "i(ee)," and prevent the tones' becoming too dull or blunt.

The frequency location of the main formant, as also the intensity of the secondary formants, depend within certain limits on the performance technique. If a higher intonation is required it can be achieved by raising lip pressure. The drop of sound level associated with that must be compensated by a simultaneous increase in air pressure. While the sound level remains the same, the formant rises slightly along with the upward pressed tone, and the overtones gain in intensity (Smith and Mercer, 1973). The tonal character as a whole, thus becomes brighter. This effect must be considered when choosing the reed: the lower the reed is tuned, the brighter will be the tone color, assuming the same intonation.

3.2.4.2 Dynamics

The brightening effect of the secondary formants is most noticeable for mid-loudness levels. For *ff* the overtone series reaches to above 12,000 Hz, so that these very high tone contributions largely determine the tone, and the maxima around 2,000 and 3,500 decrease in importance. Here, however, directional effects also are relevant (see Sect. 4.3.4), since the power level spectrum decreases above 1,000 Hz with a slope of 20 dB/octave. The high frequency overtones of the lower registers are distributed so densely that they assume a noise-like character, which can lead to a certain hardness when the tone is too forced. Additionally, the secondary maximum, which for *mf* lies near 1,150 Hz, and is shifted into the region of the nasal components for *ff*, exhibits a greater intensity in the upper registers than the principal maximum. Through this the tone loses sonority.

In contrast, at lower loudness levels, the principal formant is shifted toward a darker tone color, furthermore, for *pp* the power spectrum above approximately 600 Hz already drops by 25 dB/octave in the low register, and by 35 dB/octave in the middle and higher registers, so that the tone becomes rounder and more damped. However, for *pp,* noise contributions in the frequency region around 3,000 Hz can be formed in the middle register, which come about because of Eigenresonances of the reed. They depend in their intensity within wide limits on reed properties (Meyer, 1966c). The consequence of this is that the dynamic tone factor of the bassoon – at least in the lowest octaves of the range develops tendencies which are

contrary to usual expectations: When the strongest tone contributions in the region of the main formant are changed by 1 dB, the components in the region around 3,000 Hz are strengthened or respectively weakened by 0.6–0.9 dB, only from the third octave upward does the "dynamic tone color factor" follow the usual tendency with values of 1.2–1.5.

In the bassoon, the dynamic range depends in particular measure on the speed of performed notes. For rapid scales the power level can be varied between 81 and 96 dB; in the mid register long notes can be weakened to 72 dB and raised to 102 dB, which corresponds to a dynamic range of 30 dB. For the low and high registers this is narrowed to a dynamic range of about 25 dB. The power level for a mid- range *forte* lies near 93 dB, as for the oboe and the clarinet.

3.2.4.3 Time Structure

In spite of the low pitch of the instrument, the attack of the bassoon is very precise. The reason for this lies in the fact that the overtones in the middle and high frequency regions have a very short initial transient; they already reach their final strength within about 20 ms as already shown in the example in Fig. 2.2. Through this, the beginning of the tone is clearly defined. As is the case for other reed instruments, no additional noise accents appear, so that the attack gives a very pure impression. Nevertheless, as seen in Fig. 3.13, the low frequency contributions (below about 200 Hz) for such sharply attacked notes require a longer time for the initial transient – a value of 20 ms would only involve about two vibrational periods for the fundamentals of the lowest notes.

Since the lower tone contributions are relatively weak from an intensity standpoint, their initial transient time of 50–80 ms does not reduce the precision of the attack, however, a short staccato in the lowest registers can influence the sonority. For the bassoon, most references in the literature give an overall value, without frequency weighting, for the initial transient time in the order of 30–40 ms, which characterizes the precision of the attack in comparison to other instruments (Luce
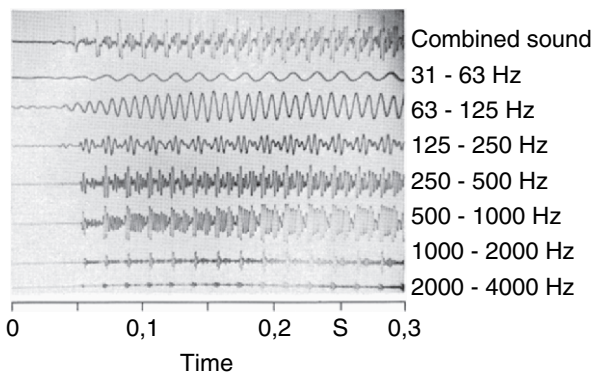

Fig. 3.13 Octave filter oscillogram of the initial transient for a bassoon (played note B_1^b)

Score example 8 C. M. von Weber, Der Freischütz, Beginning of the overture

and Clark, 1965; Melka, 1970). Even for a soft attack, the initial transient is not lengthened significantly above 60 ms, however, the tonal impression in that case is primarily determined by the fact that the higher frequency components enter later than the lower components and that the initial transient is slower than for sharply attacked notes. This difficulty, to achieve a soft entrance, is noticeable, for example, at the beginning of the Freischütz overture, when the bassoon is expected to merge unnoticeably into the "from Nothing" swelling string sound, without lending an accent to the joint beginning. (Score example 8).

The decay processes are short, as is the case for all wind instruments. The decay time lies at around 0.1 s for the higher notes (as was recognized from Fig. 2.2) and can be lengthened to 0.4 s for the lowest registers (Rakowski, 1967). The influence of the player is minor.

When bassoon players play a pronounced *vibrato*, the width is about ±15 cent. This results in level fluctuations of only 4 dB for the strongest partials, i.e., in the region around 500 Hz. The strongest level fluctuations occur between around 1,400 and 1,800 Hz and can reach up to 15–20 dB. Since, by reason of their time structure, they draw the attention of the listener, they emphasize the nasal quality of the sound. While all overtones fluctuate in phase, one could speak of a tone color modulation, still, the tone receives an inner stability through a number of weakly modulated partials with fluctuations of only 5–6 dB, which do not permit that effect to become as noticeable as is the case with the brass instruments and the flute.

3.2.4.4 The French Bassoon

While the bassoon of so called German construction is preferred in almost all countries, in France and also in a few east European countries instruments are played with different fingering and different tonal characteristics. These so called French Bassoons were historically developed alongside the German Bassoon and represent a certain parallel to the Horns in France. In the tone picture, the nasal components are also strongly pronounced, as the juxtaposition of the formants shows in Fig. 3.14. While the basic formant in comparison to the German instruments is shifted in the direction of a darker "o(oh)," all secondary formants, however, are located noticeably higher. As a whole, the tone of the French Bassoon is richer in overtones, and the vowel color of the basic formant (with a logarithmic decrement of 1.6) is somewhat less firmly determined, since it is broader than in the German model.

By reason of all these characteristics, French bassoons sound less forceful, and have a more slender tone, which in connection with the good attack at the pitch of these instruments is particularly suited for virtuoso passages. In their tonal character, therefore, they serve the compositions of the French impressionists particularly well, as also the older, preferably virtuoso wind literature, while the German bassoons are based more in the tonal conception of the romantic epoch.
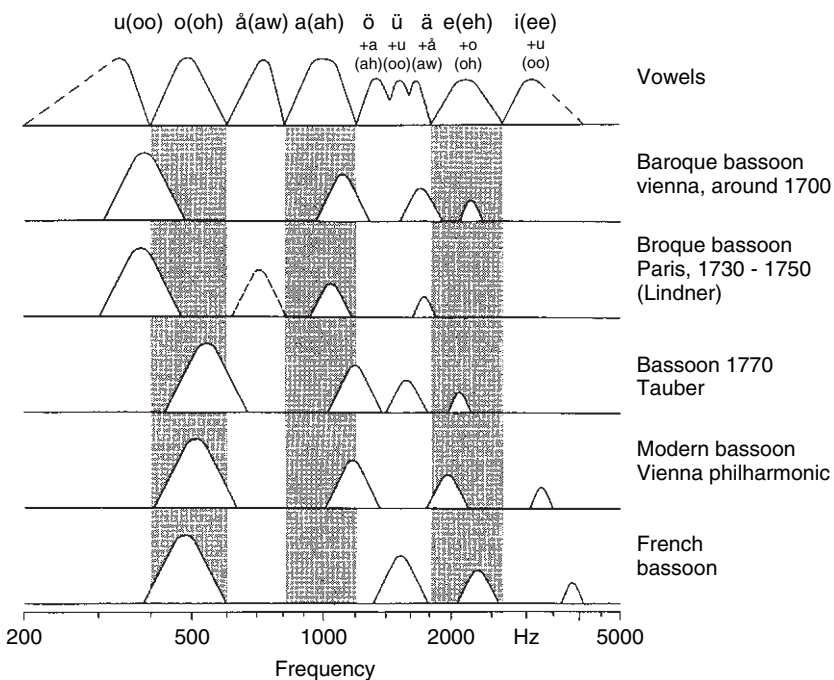

Fig. 3.14 Schematic representation of the formant location of several bassoons

3.2.4.5   Historic Bassoons

While data about instruments in use today are based on a large number of instruments investigated, and thus are generally applicable, for historic instruments, naturally only individual examples can be cited. For the bassoon, results are available for two baroque instruments, and one example from the classic period, all three instruments are well preserved. They are still being played in ensembles for old music. While they are not as balanced as today's instruments, some typical characteristics can be drawn from their formant locations (see Fig. 3.14) (Meyer, 1968).

Particularly noticeable is the low frequency position of the main maximum for Baroque bassoons, which results roughly in a bright "u(oo)" as the basic color. This, combined with the lower location of the secondary formants and the already relatively small overtone content, brings about a tonal effect darker than for today's instruments; it can on occasion even become somewhat muffled. To this is added, that the logarithmic decrement of the main formant in the lower octaves of the tonal range lies at around 1.2, while in contrast, it is above 2.0 for the upper registers. This signifies, that the tonally best regions for the Baroque bassoon are located in the lower regions; thus, these instruments seem especially suited for supporting the continuo foundations or for characterizing sinister moods in operas or oratorio scenes, as for example the appearance of Samuel's shadow in Händel's "Saul."

In contrast, the tonal picture of the bassoon looks entirely different in Mozart's time. The main formant is situated in the region of the vowel "o(oh)" as with today's instruments, however, for the upper locations of the tonal range, it moves right up to the boundary with "å(aw)." The secondary formants are concentrated near a very bright "a(ah)" nasal component, and finally a middle "e(eh)," so that they add a character which is not too weak but slender, to the already relatively bright fundamental color. The logarithmic decrement of the main formant with a value of 2.1 in the lowest register, is less favorable than for the higher notes with values of 1.4–1.6. The best location from a tonal standpoint, i.e., the region with the most pronounced formants, is therefore in the middle and the upper registers, while the low register is somewhat duller. With that type of timbre, the instrument is particularly suited for supporting precision in the bass groups as also in higher song-like passages.

3.2.4.6   The Contrabassoon

With a lower limit of the tonal range at B_0^b or even A_0 (e.g., in "Salome") the contrabassoon, next to the contrabass tuba, is the lowest instrument in the orchestra. Its spectral region thus begins at 27.5 Hz. However, the radiated energy at that pitch is very low. Below the first maximum in the power spectrum the level drops – as also for the bassoon – with 8 dB/octave. This first maximum in the overtone structure is located at about 250 Hz. Secondary formants are found near 400–500 Hz as well as in the region around 800 Hz for greater loudness (Meyer, 1966c). The

basic color can, therefore, be described as a dark "u(oo)" with an addition in the region of a dark "o(oh)."

Because of the double bend in the body, the tonal development over the range of the instrument is not as uniform as for the normal bassoon. Particularly in the lowest register the strongest tone contributions switch between partials of different order. Since the ear is significantly more sensitive in the region around 400 Hz than in the region of the lowest partials, the strongest partials in the sensitivity spectrum become more prominent and make the pitch orientation more difficult for the lowest notes. For example, in the descending tone sequence B_0– B_0^b the rise between the strongest partials, namely from about 395 Hz (13th partial of B_0) to about 405 Hz (14th partial of B_0^b) can be heard very clearly. This results in a perceived uncertainty about whether the note sequence rises of falls.

Noteworthy, however, is the relatively short initial transient for the contrabassoon. For tongued notes in all registers this amounts to only from 30 to 35 ms, so that the instrument gives an impression of rather high agility in spite of its low pitch (Melka, 1970). On the other hand, its dynamic range in comparison to a bass clarinet, is relatively narrow: for individual notes it does go up to 15 dB, yet in a melodic context it drops to 10 dB. This is caused by a relatively high lower limit of the dynamic range with a power level of 86 dB at a *pp,* while for *ff* the 96 dB value comes close to the upper limit of a normal bassoon.

3.3   String Instruments

3.3.1   The Violin

3.3.1.1   Sound Spectra

For wind instruments the tonal picture is primarily determined by resonance effects of the enclosed air column, the vibrational characteristics of which are determined by the dimensions of the bore diameter, the bell, the mouthpiece etc; the material, in contrast only plays a subordinate role. For string instruments, on the other hand, the tone is primarily formed by the resonance characteristics of the corpus. Since, however, the acoustical properties of the wood differ from instrument to instrument, the result is a multitude of variants.

The basic structure of the power spectrum of string instruments is determined by the spectrum of the string vibrations excited by the bow. In this spectrum the fundamental dominates, and the subsequent overtones decrease with increasing frequency at a rate of 6 dB/octave. Above the bridge resonance, i.e., above approximately 3,000 Hz for violins, the power level decreases by 15 dB/octave. Superimposed on this is the shaping of the tone by the corpus, which possesses a large number of resonance regions, fixed in frequency, they thus are not shifted with performed pitch – as is the case for wind instruments. For a chromatic scale, therefore, the excitation spectrum moves over the resonance chain, which remains

fixed in frequency, so that the tonal character can change from note to note. The tonal spectra of string instruments are thus not as uniform and systematically structured as those of wind instruments, but rather exhibit a greater range of color (Leipp, 1965; Lottermoser and Meyer, 1968).

By reason of this cooperation between string vibration and resonance body, the fundamental is almost without exception the strongest partial of the spectrum for the violin in the upper registers. This is valid for the entire range of the E-string, for the A-string from about E_5, and for the D-string from about C_5 on up. In addition, a dominating fundamental appears in two especially well defined resonance regions of the violin, namely around G_4 and between C_4 and D_4. This lowest resonance is associated with the normal mode of the air volume enclosed in the body and is designated as the air resonance. Below this resonance, the fundamental drops in the power spectrum by about 4 dB per half step, so that it lies around 20–25 dB below the strongest partial on the G-string.

As an example of sound spectra of the violin, Fig. 3.15 gives measurement results for the four lowest notes of an instrument with very good tonal qualities. In comparison with the chromatic sequence for a horn in Fig. 2.4, the greater multiplicity in the distribution of partial intensities in the violin is clearly recognized. Thus, neither a steady course of the envelope nor the domination of odd partials can
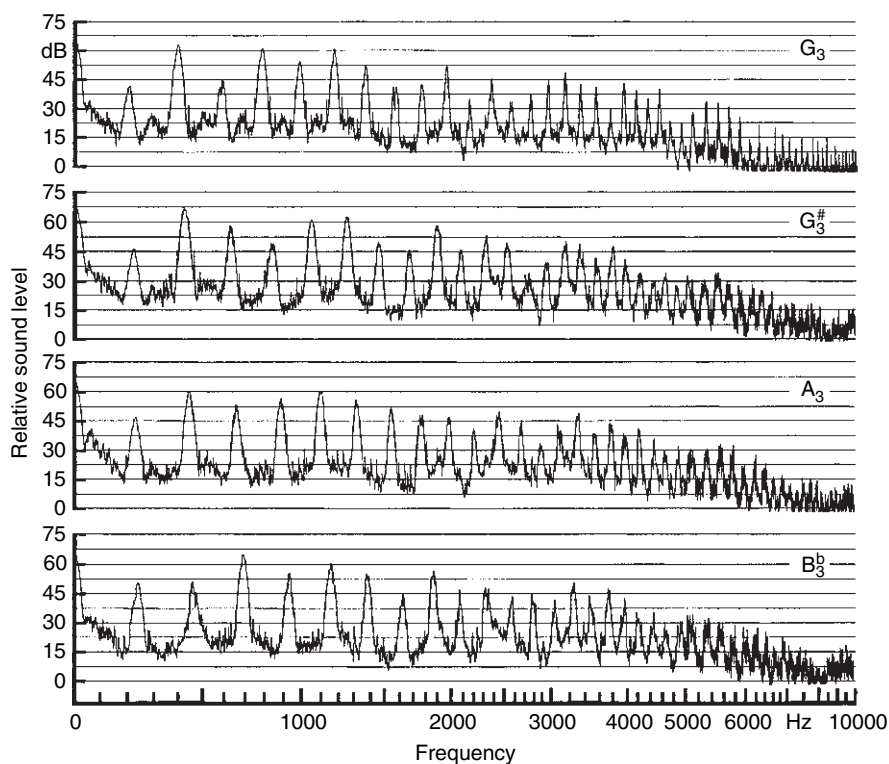

Fig. 3.15 Sound Spectra of a violin by Guarnerius del Gesù (1739)

be read from the curve, rather the spectra are characterized by individual prominent partials which happen to coincide with a resonance of the corpus. For G, G$^{\#}$, and A, pronounced octaves follow weak fundamentals, while for B^{b} the third partial is strongest. Beyond that, up to nearly 10,000 Hz the spectra then show individual partial groups of larger intensity which show a formant-like character. For example, this occurs between 1,000 and 1,200 Hz as well as between also between 3,000 and 4,000 Hz.

As evident, the formant location for the individual notes, varies somewhat, it is entirely possible that for certain notes, in relation to the location of body resonances, no such predominant partial groups are formed. The number of notes within the full chromatic range, for which formants appear, is therefore a typical quality criterion for the instrument, and the vowel-like tone color supports the song-like tonal character. For example, in the well known Stradivarius violin "Prince Klevenhüller" for 40 of a total of 52 well studied notes, formants could be demonstrated (Lottermoser and Meyer, 1968). In this context it needs to be emphasized that the production of these strong partial groups, in contrast to other tonal characteristics, depends only in very small measure on the quality of the performance.

Typical violin formant locations for the low notes are around 400 Hz; as a dark "o(oh)" these sounds provide the sonority for the lower G-string. In the region from 300 to 350 Hz, old Italian violins of the Stradivarius type already show a higher intensity than French violins or most newer instruments (Meyer, 1982a; Dünnwald, 1988). Through this, the tonal difference between the notes in the region of the air-resonance (about B$_3$ to D$_4$) and the lower notes of the D-string is especially diminished. A second formant region, covered by the vowel color "a(ah)," i.e., between 800 and 1,200 Hz, is very essential. This partial group must be considered as particularly characteristic for the violin sound, it lends strength and substance, and prevents the nasal timbre. The exact frequency location of this formant is very significant for individual traits within the tonal picture: in violins of Guarnerius del Gesù tonal contributions between 1,000 and 1,250 Hz are almost always especially emphasized, as also noted in Fig. 3.15. These instruments are therefore even stronger than the Stradivarius violins, for which the maximum clearly lies below 1,000 Hz. Instruments with a pronounced dark timbre have the corresponding formant even below 1,000 Hz, that is at a tone color tending more toward an "å(aw)"; This characteristic is also found in several Stradivarius violins. In this context it should be noted that a spectrum with a del Gesù-like sound, with maxima near 400 and 1,200 Hz, was perceived as particularly "well-sounding" in subjective hearing tests among a multitude of tones. (Terhardt and Stoll, 1978).

The next higher frequency region around 1,600 Hz, which is responsible for the covered tone coloring and also for a nasal timbre, is radiated more strongly by a decided majority of old Italian violins in contrast to other instruments. This removes a certain hardness or directness from the sound. A series of further formant regions are responsible for the brilliance and brightness of especially the upper tones, the most important of these are between 2,000 and 2,600 Hz, i.e., in the vowel region of "e(eh)," and between 3,000 and 4,000 Hz in the region of the vowel "i(ee)."
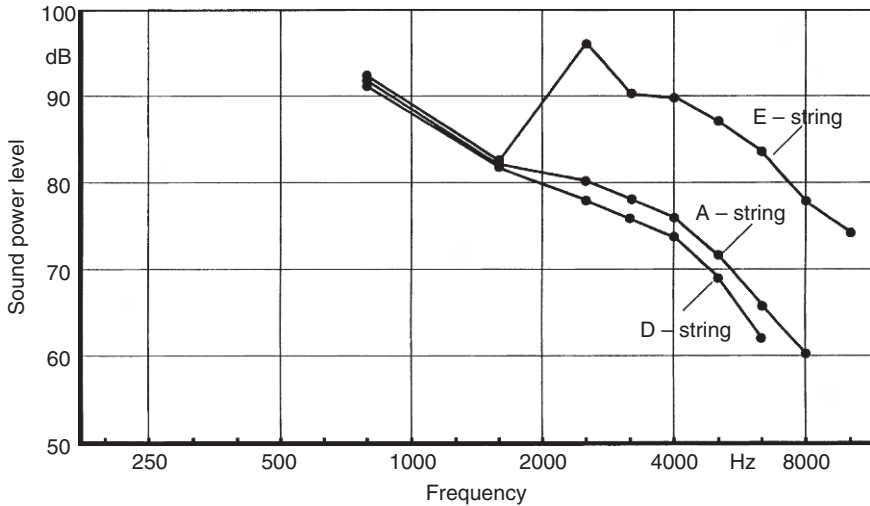
Fig. 3.16 Power level spectra of a violin, the note A₅ is played *ff* on three different strings (after Meyer and Angster, 1981)

The frequency range of violin tones depends on various components of performance technique. The largest number of overtones appear with an open string, since the string termination is more sharply defined at the nut than is possible with a pressing finger. However, the more firmly the finger is pressed, the richer in overtones, and the more precise other notes become. Naturally the relative position of the note also is of importance. Thus, the change from the D-string to the G-string shows two characteristics typical for the transition to a darker tone color: the formant maxima shift to somewhat lower frequencies, where especially the region around 1,000 Hz is weakened, furthermore, the overtone content above about 2,000 Hz decreases strongly. While the difference between D and A strings is least pronounced, the overtone content increases significantly in the transition from the A to the E string. Fig. 3.16 shows a comparison of power spectra for A₅ played on the D –, A – and E – strings with equal strength. Above about 2,500 Hz, i.e., above the bridge resonance, the overtone level of the E – string lies about 15 dB above the corresponding tone contributions of the A – string, in contrast the D – string has an overtone radiation of only about 3 dB less than the A – string.

The sound spectrum depends strongly on bowing technique. The three influential quantities to be modified by the performer are: the speed of the bow, the pressure exerted by the bow hair on the string, as well as the point of contact at which the bow touches the string. The bowing speed influences fundamental and overtones equally, it is, therefore, the most important means of influencing dynamics. In contrast, bow pressure has no influence on the fundamental, assuming equal bowing speed, an increase in bow pressure primarily raises the intensity of higher overtones (Bradley, 1976; Cremer, 1981).
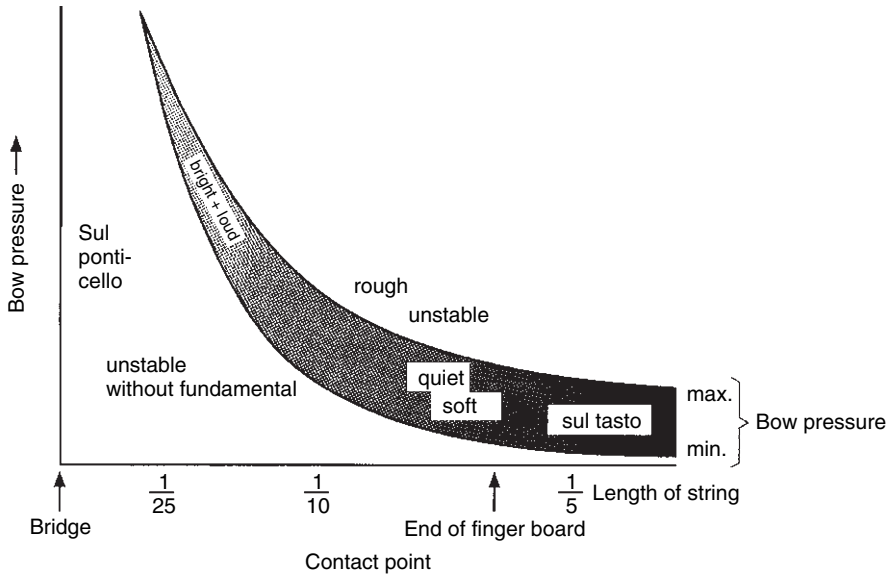
Fig. 3.17 Dependence of bow pressure on contact point for constant bow speed, The shaded area represents the actual performance region with its tonal and dynamic expression possibilities

    The point of contact, on the other hand, influences the entire spectrum. Thus, the closer the contact point is to the bridge, the more bow pressure is needed. In Fig. 3.17 this relationship is represented. Below the line of minimal bow pressure, the available energy is insufficient to maintain a stable vibration. Above the line of maximal bow pressure, the restoring force of the string is insufficient to force strictly periodic vibrations. Increased fluctuations and added noise contributions, therefore, cause the tonal picture to be rough, which can create the impression of scratching. The closer the contact point lies to the bridge, the brighter and louder the sound becomes. However, the variational range of the permissible bow pressure also becomes narrower. Soft sounds with covered coloration arise from a contact point near the finger board (Schelleng, 1973; Meyer, 1978b).

    The attack noise, generated by the bow, in addition to the harmonic partial spectrum, belongs to the typical tone characteristics of string instruments (Lottermoser and Meyer, 1961). The fact that this hissing noise is a specific component of the tonal picture has become particularly evident in experiments with electronic imitation, where the harmonic spectrum alone could not create the impression of a string instrument. As already suggested in Fig. 2.6, the bow noise receives a typical coloring through the body resonances. This is the same for all notes of an instrument, and it stands out, especially in hearing impressions in the upper registers, since here the partial spacing is greater. Irregularly excited vibrations of the air volume (around C_4) of unique pitch are audible as a color noise, and so is a broader resonance region above 400 Hz. In comparison to wind instruments, these noise contributions – particularly for lower dynamic levels – are relatively strong

for strings. For equal loudness of the harmonic tone contributions, the noise components of string instruments are by about 20–30 dB stronger than for wind instruments, with the exception of the flute.

The shaded area represents the actual playing region with its tonal and dynamic expression possibilities.

3.3.1.2 Dynamics

Dynamics for string instruments are essentially responsive to bow speed and the location of the contact point. Assuming that the bow speed, in practical terms, can vary between 10 and 125 cm s^{-1}, this corresponds to a dynamic range of about 22 dB. The distance of the contact point from the bridge, between the closest and most distant point increases by a factor of six, which corresponds to an additional 16 dB. In principle, therefore, one can count on a dynamic range of just under 40 dB. This range, however, cannot be fully exploited. For some notes, strong resonances make a *forte* easier, they also make a *pp* very difficult. Thus, near the air resonance it becomes difficult to make an extreme *pp* speak.

The softest *pp* which can be performed on a violin depends on the strength of the basic noise. This, in turn, maintains a fixed intensity from *pp* to beyond *mf* and only increases for very large playing volumes (Lottermoser and Meyer, 1961). The lower limit of the dynamic range of the violin is therefore given by the condition, that the partials must extend sufficiently clearly beyond the noise. As a result, some favored isolated notes have a *pp* power level of only 58 dB, however, this rises to 74 dB for rapid tone sequences. For *ff* a violin reaches power levels of 94 dB for fast sequences and 99 dB for slow ones, so that one can count on a practical dynamic range of about 30–35 dB. The power level of a median *forte* lies near 89 dB. (Meyer, 1990).

3.3.1.3 Time Structure

The excitation of string vibrations by the bow offers far greater flexibility for shaping the initial transient than is the case for most wind instruments. Nevertheless, even for sharp attacks on a violin, initial transient times are longer than for example on an oboe. On a G string, a fast initial transient process lasts almost 60 ms (see also the upper example in Fig. 2.2), this, however, is shortened in the upper registers. In the mid-range values between 40 and 50 ms can be expected, on the E string the initial transient time is shortened down to almost 30 ms (Melka, 1970). For gentle attacks the duration of the initial transient process can be extended to 200–300 ms without having the tone disrupted in its slow development. Thus a very rounded tone picture emerges.

Inasmuch as the strength of the higher overtones depends essentially on bow pressure, their development is somewhat delayed, which does not aid a *staccato* attack. Deliberately high bow pressure in an attack (e.g., in a *detaché*-stroke) creates a broad-band articulation noise, with a duration of about 50–100 ms.

Since this noise contains preferentially the frequencies of the strongest resonances, the vibrations of these resonances can be heard within the total sound, partly with an almost tonal effect, which leads to a basic color of the tone, common to all sounds of the instrument. These initial transient effects of the resonances are most noticeable while playing *col legno*, because in that case the string vibrations do not come to full development.

Finally, it also needs to be mentioned, that during an attack or string change, the frequency of the note played is lowered by 10–20 cents when compared to the final value. This also influences the aural impression of the total "initial transient" phenomenon, even if a pitch variation does not become directly noticeable. What is perceived is merely the less precise attack when compared to an oboe or clarinet.

The decay time for string instruments depends in large measure on whether the bow remains on the string or is lifted. As the middle example of Fig. 2.2 has shown, the decay time in the first case is only about 0.1 s. If the bow is lifted, the tone will continue to sound longer, depending on the length and mass of the string. For a violin, decay times of 1 s in the lower and 0.5 s in the upper registers are typical. For open strings, decay times lie between 2 and 3 s. All these values are given in relation to the relevant fundamental; the first overtones continue to vibrate at most half that long, and the higher overtones are practically damped immediately (decay times of less than 50 ms).

A pronounced *vibrato* for violinists has a width of the order of ± 30–35 cent. These frequency fluctuations of the played note lead to the circumstance that the individual overtones in their increasing and decreasing frequencies move up and down on the flank of corpus resonances and are modulated in intensity. Depending on whether they fall on the rising or falling flank of the resonance their intensity pulsations occur in equal or opposite phase. This effect is clearly recognizable in
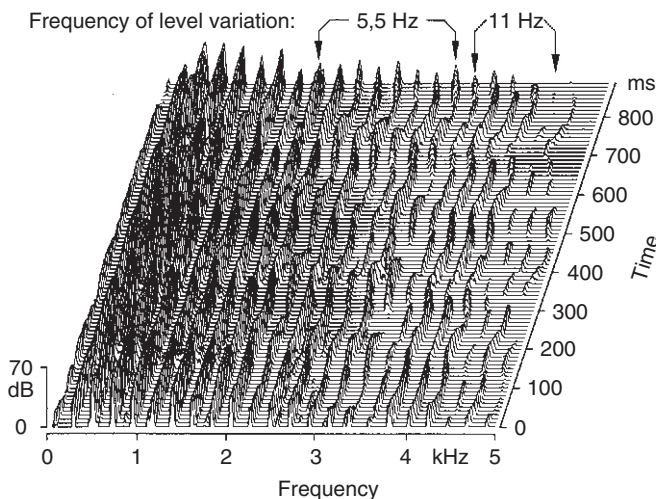

Fig. 3.18 Changes of the tonal spectrum with time for a *vibrato* in bowed strings (Violin, note A_3)

Fig. 3.18: The 8th partial (at about 1,800 Hz) has its maximum at the frequency
maximum, while, for example, the 15th partial (at about 3,300 Hz) has its maximum
at the frequency minimum. It can also happen that the frequency of a partial
straddles a resonance peak, and that thus during a period of the frequency motion
two amplitude maxima and minima appear. An example of this amplitude modula-
tion at twice the frequency is the 20th partial (at about 4,400 Hz) while the 16th
partial (at about 3,500 Hz) runs through a low point, and thus correspondingly
pulsates with a phase opposite to that of the 20th partial.

When a pronounced higher partial is modulated in this fashion at twice the
frequency, it can stand out in a penetrating way and lead to a kind of roughness of
the sound. While in exceptional cases individual partials can exhibit fluctuations of
up to 25 dB, in most cases for frequency groups relevant to the ear, a partial
compensation between components of opposite phase occurs, so that the ear only
senses fluctuations of the order of 10 dB. The lower partials, still perceived by the
ear as separate, fluctuate typically by 3–6 dB below 1,000 Hz and by 6–15 dB
between 1,000 and 2,500 Hz at a vibrato of ± 35 cents (Meyer, 1992).

The time structure plays an important role in polyphonic chords, since they can
either be attacked sharply or be broadened by arpeggiation. The left picture of
Fig. 3.19 shows the temporal spectrum development for the notes G_3 (green),
E_4 (yellow) and C_5 (red). Disregarding the attack noise, one can recognize that
G_3 and E_4 enter simultaneously and that C_5 follows with a delay of about 40 ms; the
overtones are developed slightly later. After 170 ms the G_3 is no longer excited, as
is evident by the detaching overtones. For the deliberately broadly attacked *arpeg-
gio*-triad in the right picture the notes E_4 and C_5 only follow the low G_3 by about
250 ms, this note in turn decays quickly, and only continues in the fundamental
and the octave partial. These two examples provide the limits of the width of
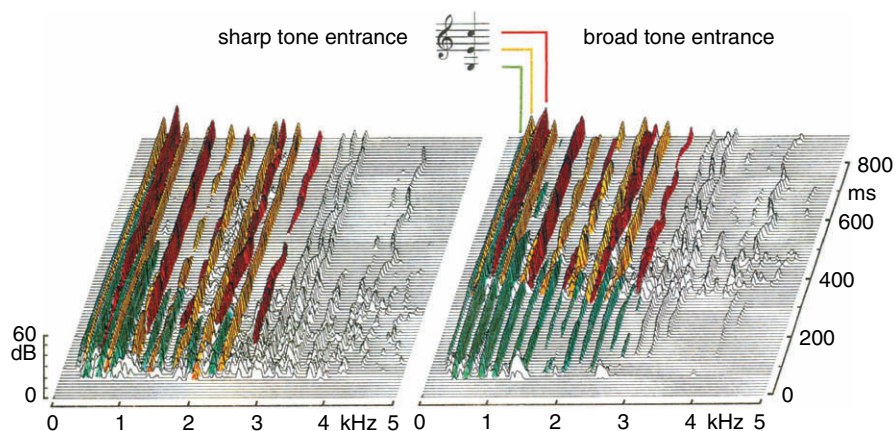possibilities for a temporal structuring of such chords.

Fig. 3.19 Time evolution of tone spectrum for a three note violin chord with two different playing
approaches (See Color Plate 1 following p. 178)

3.3.1.4 Special Performance Techniques

For a *pizzicato* the string is displaced slightly from its rest position and can vibrate freely after release. This results in an extremely short initial transient, which for the violin involves less than 10 ms in all registers (Melka, 1970). Inasmuch as the noise components, which are also excited, lie far below the actual partials in strength, the lower limit of the dynamic region can be reduced to 51 dB. For *ff,* in contrast, sound pressure levels up to 90 dB can be achieved, albeit only for short durations. Accordingly the dynamic range of 39 dB is of the same order of magnitude as for *arco* performance.

The spectrum for pizzicato tones depends largely on the plucking location. For plucks near the bridge, the sound becomes rich in overtones and hard. A softer timbre results above the fingerboard. It assumes a somewhat covered character when one plucks in the middle between the bridge and the fingering touch, since at that position the even partials are only excited weekly and the odd partials dominate.

The audible decay, and thus also the duration of *pizzicato* tones, naturally depends on the loudness of a performance. For *pp* it varies between 40 and 150 ms and for *ff* between 350 and 800 ms. Open strings have the longest decay times, and notes of equal pitch decay more rapidly when fingered in high positions in comparison to low positions (Spelda, 1968). Furthermore the upper overtones decay more rapidly than lower contributions, as is noted from Fig. 3.20. Thus, the decay time, as noted in the upper picture, is approximately 3 s for the fundamental and 1.5 s for the first overtones. The influence of the *vibrato* on the tone of the string *pizzicato* is noteworthy. As illustrated by the lower picture, the tone initially sounds more lively. It becomes notably dryer, during the *vibrato* since the motion of the fingers provides additional dampening. The decay time is thus cut in half when compared to the note without *vibrato*. In the example pictured, to 1.5 s for the fundamental and from 0.7 to 0.8 s for the subsequent overtones. This effect is unnoticed by many string players (Meyer, 1992).

Placing a mute on the bridge dampens primarily the high-frequency contributions. However, there are some additional effects, the details of which depend on the nature and the weight of the mute. With increasing mute weight, the resonance of the body, normally found at 400 Hz, is shifted to lower frequencies. It thus approaches the air resonance. This leads to an increased intensity of the fundamental for the lower notes of the D string. A determining factor in this is also the strength and the weight of the bridge, so that the effect varies from instrument to instrument. Resonance shifts can also occur in the region of higher frequencies, without, however, causing higher intensities than for normal play.

The amount of pitch-damping is important in its effect on tone color. Most mutes dampen components down to those frequencies which lie above nasal contributions. As a result, the nasal tone is created which frequently is considered typical for violins with mutes. Particularly light mutes naturally cause the least change, creating a somewhat lighter and barely nasal tone. They are especially suited for
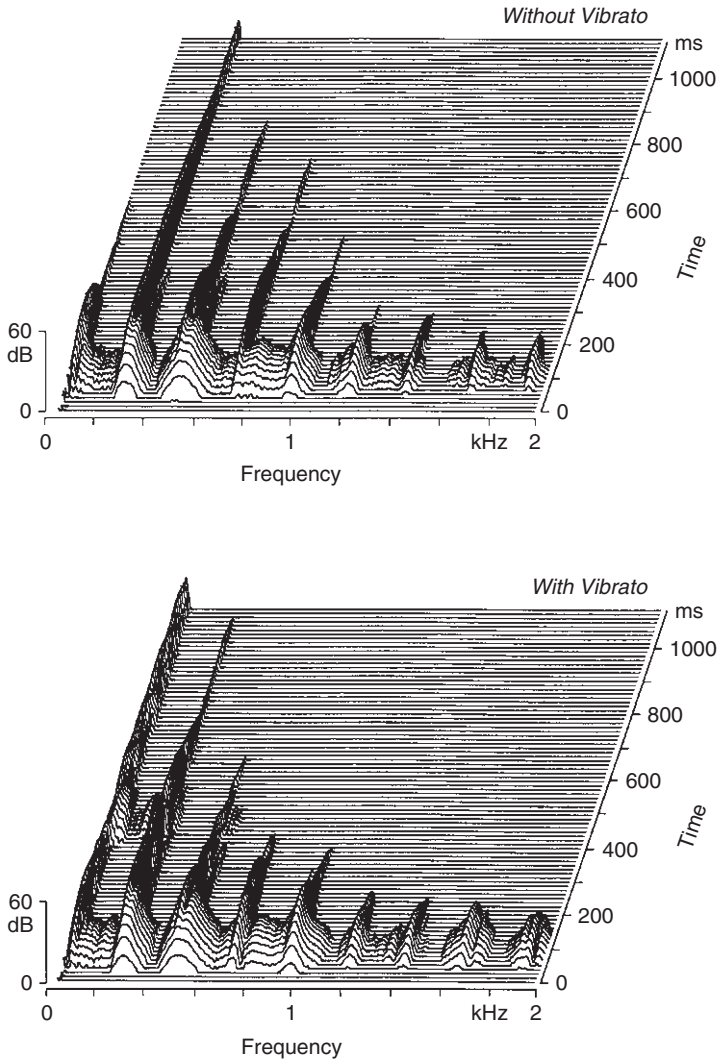
Fig. 3.20 Influence of *vibrato* on time evolution of the tonal spectrum of a string *pizzicato* (Violin, B$_3$)

low passages, for which the resonance shifts, mentioned earlier, toward the lower notes of the D string caused by heavy mutes would be disturbing. On the other hand this type results in a very soft and covered sound for the A and E strings, since they also reduce the nasal contributions. However, this does require a weight of about 6 g which is available in a five prong steel mute, for example. Through the choice of suitable mutes, a number of different tonal variations can be achieved. For most mutes a reduction of radiated sound power of the order of 6 dB results. For very heavy mutes this can amount to 10 dB.

"Harmonics" (also called *flageolet* tones) are played by gently touching the string. The fundamental of the open string is also excited very weakly by the bowing noise. The noise background thus receives a character similar to the flute, resulting in the unique timbre of such tones. Since the excitation for these tones is rather critical, the initial transient is somewhat longer than for normally fingered notes, however, because of the relatively large energy of the entire string vibration, the decay time is quite large. The dynamic range is around 20 dB, significantly narrower than for normal tones. It is possible to play a very gentle *pp* with a power level of 64 dB, however, the upper limit of an attainable power level of 84 dB is very low (recalculated from Burghauser and Spelda, 1971). For so called "artificial harmonics," which are created above a fingered note, the dynamic range is yet more stringently reduced.

3.3.2 The Viola

3.3.2.1 Sound Spectra

Tonal characteristics of the viola naturally resemble those of violins in essential features (Fletcher et al., 1965). Lower tuning (by a fifth) of the viola extends the range down to C_3 (130 Hz). This means that for the lowest radiated frequency the wavelength is 1½ times larger than for the violin. The dimensions of the viola, depending on body size, however, are only from 1.15 to 1.2 times larger than for violins. Thus, the typical resonances are lowered by significantly less than a fifth when compared to a violin, and the tone color is not that much darker, which would correspond to a tuning lowered by a fifth. Therefore, many attempts have been made to build larger violas in an effort to increase the fullness of sound in the low register.

The air resonance of violas generally lies in the neighborhood of A_3, around 220 Hz. Correspondingly, the amplitudes of the fundamental of the C string are very weak. For C_3 they can lie by more than 25 dB below the strongest partial. The low notes of the viola have their largest intensity in the two resonance regions around 220 and 350 Hz. These two fall into the vowel region of "u(oo)." They provide substance for the tone. Further formant regions, which primarily affect the mid-register and high notes are located around 600 Hz, which is near the transition between the vowel colors of "o(oh)" and "å(aw)," as well as around 1,600 Hz. The last maximum brings out the frequently noticed nasal timbre of this instrument group. As a whole, the intensity of the higher overtones is less than for violins. The resulting tone is therefore not as brilliant. In good violas, however, the timbre is additionally brightened by a secondary formant in the region of the vowel color "i(ee)" (between 3,000 and 3,500 Hz). This also weakens the nasal effect.

3.3.2.2 Dynamics

The lower limit of the usable loudness region for the viola lies slightly above that of the violin: In the low registers slow notes can go down to a power level of about

67 dB at *pp*, and in the regions of the D and A strings to about 63 dB. For a *ff*, in contrast, the radiated power level is somewhat lower than for the violin. It lies at about 95 dB, and only the rarely used pitch range above C_6 drops to values below 88 dB. For rapid note sequences the usable dynamic range drops to a region of 73–91 dB. Accordingly, the dynamic range of the viola is slightly less than that of the violin, which is also related to the somewhat more sluggish initial transients in the extreme registers. As with other bowed string instruments, the influence of performance technique on the higher overtones is small. The "dynamic tone color factor" lies near 1.1. With a value of 87 dB the power level for a median *forte* is by 2 dB lower than for violins.

Only when playing harmonics does the viola exceed violin radiation by about 3 dB. The reduction of radiation by using a mute generally does not reach the same values as for violins, since mutes in relation to bridge mass are not as heavy. A level reduction of about 4 dB can be expected.

3.3.2.3   Time Structure

In spite of its larger dimensions, for sharply attacked notes the viola does not require a longer time for full tone development when compared to the violin, in fact, frequently even shorter initial transients were noted. Accordingly, initial transients for *staccato* notes occupy about 30 ms. These values are valid even for the lowest registers, where the open C string can exhibit an initial transient of 20 ms (Melka, 1970). This result not only requires optimal conditions for string and bow, but it is also deceptive in as much as it covers the fact that the already weaker lower tonal contributions do not reach their final intensity in such a short time. This is hardly annoying for long notes. However, for very short notes the tonal character of the instrument is shifted toward a brighter, and often nasal tone color.

The possibilities for expressive shaping of an attack are not equal for all pitch ranges of the instrument. Thus, in the range of the C string the initial transient can be broadened only to 100 ms by an appropriately soft attack. With rising pitch, the variation possibilities increase. On the A string, tone development times of more than 200 ms can be achieved in the upper registers. In this context it is noteworthy that these soft attack values for the viola are also less than for the violin, and that they exhibit a tendency opposite to the violin: in a violin the lower notes are more strongly influenced than the high ones.

For the viola, the decay time is negligibly longer than for the violin. For detached bowing, values around 1 s are observed for fingered notes in the mid-register, and 3–4 s for open strings. As is the case for violins, overtones decay more rapidly than the fundamental.

3.3.2.4   Pizzicato

The initial transients of plucked strings on a viola are nearly as short as for the violin. Only on the C string is this transient slightly longer than 10 ms, for higher

strings this time is not exceeded (Melka, 1970). However, violin intensities are not achieved. For *ff* the power level is about 90 dB, and for chords over all four strings this value can be raised by 3 dB. The softest *pp* can be assumed to be about 58 dB, so that even in regards to the lower limit a narrowing of the dynamic range results, when compared to the violin.

The decay of *pp* tones lasts 50–150 ms, for *ff, pizzicato* tones reach a duration of 280–600 ms (Spelda, 1968). These values are similar to those for violins. A certain difference comes from the fact that with the viola the open strings decay noticeably longer (500 ms on the average) than fingered notes (about 400 ms), while this difference is not as pronounced for violins.

3.3.2.5   Viola d'amore

Even though the Viola d'amore blossomed already in the time of the Baroque and the early classic period, this instrument, with its peculiar tone color is found occasionally in modern operas such as "Katja Kabanova" or "The Affair Makropoulos" by L. Janacek, or "Louise" by G. Charpentier. These examples follow H. Pfitzner, who used the instrument in his "Palestrina," and G. Puccini, who used its peculiar sound effect as an accompaniment of a choir behind the stage in "Madame Butterfly."

The large tonal range from A (110 Hz) up to the height of the 6th octave, as well as the drone strings constitute the essential difference of the seven-string viola d'amore, in contrast to other bowed string instruments. When the instrument is played, the drone strings respond with their fundamental frequency or overtones and thereby emphasize various partials in the tonal picture. When – as customary – these strings are tuned to a D-major or d-minor chord, the tonic and dominant of this key receive particular brilliance, however, sympathetic resonances are found also for other keys. If, on the other hand, the drone strings are tuned to a group of neighboring half tone steps – e.g., from E_4^b to B_4^b without A_4 (Stumpf, 1970), a formant-like tonal effect is achieved in the region of the vowel color "o(oh)," however, the purity of the tonal picture is easily compromised by beats of decaying drone string vibrations. Good instruments already have a strong body resonance near F_4, which effects a tone coloring in the direction of the vowel color "o." In the upper register two additional formants are added near 650 Hz and 1,000 Hz, which shift the basic tone color toward "a(ah)," while in the low register an air resonance between 210 and 249 Hz leads to a timbre similar to a viola.

3.3.3   The Cello

3.3.3.1   Spectra

With a lowest note of C_2 (65 Hz) in the tonal range of the Cello, it is located an octave below the Viola. This difference in tone location also corresponds quite closely to the frequency distribution of the main resonances, and thus to the basic tone color of the cello sound. The air resonance for most instruments is located near

110 Hz (A_2). Below this resonance the radiated power level drops at a rate of 6 dB/octave. The lower contributions, particularly the fundamentals of the C-string, therefore have lower intensities. They can be up to 12 dB weaker than the strongest partials, when they fall on pronounced resonances as is the case, for example, for $F_2^\#$. The limited radiation of very low frequencies is especially apparent when the lowest string is tuned to B_1^b, as required, for example, in the slow movement of the R. Schumann piano quartet.

The richness and sonority of good instruments is achieved by two formant regions, which are located around 250 Hz and between 300 and 500 Hz. This causes a vowel color between "u(oo)" and "o(oh)" of the lower two strings. In the upper registers, an additional partial group gains in significance. Their maxima, depending on the character of individual cellos, lie between 600 and 900 Hz, and thus fall into the color region between "å(aw)" and a dark "a(ah)." In the frequency region of the bright a(ah)-formant, that is between about 1,000 and 1,200 Hz, the sound spectra of cellos have a pronounced depression, which is followed by the bridge resonance around 2,000 Hz.

This resonance structure of the cello body leads to the circumstance that in the region between about 200 and 2,000 Hz, a pronounced wavy nature is superimposed on the approximately 6 dB/octave drop of the power spectrum, as expected from the excitation. This causes the power level to fluctuate by ±5 dB from the steady 6 dB drop/octave (see also Fig. 7.16). This wave nature is more strongly pronounced than in the violin and therefore leads to greater tone color differences between notes separated by about a fourth or a fifth. Above about 2,000 Hz, the spectra of the lower or middle registers drop with a slope of about 16 dB/octave. In the upper register this drop is more shallow with a value of 10 dB/octave, so that the tones in the whole can become very rich in overtones.

3.3.3.2   Dynamics

By reason of its size, a cello certainly can radiate more sound energy than a violin, however, at higher frequencies it becomes clearly noticeable that the relatively large front and back plates vibrate in highly subdivided patterns, where neighboring sections vibrate out of phase, causing an acoustic short circuit. The overall sound radiation is thus of similar order of magnitude as for a violin. For rapid scales over a wide pitch range, the power level can be varied between 74 and 96 dB, for slow individual notes, between 63 and 98 dB. Above G_4 the *ff* levels, however, move around 90 dB. Thus the practically useable dynamic range covers from 25 to 30 dB. The median *forte,* with a power level of 90 dB lies by 1 dB above the violin. The "dynamic tone color factor" again lies for most notes around 1.1. When the strongest partials fall on pronounced resonances this is reduced to 1.03. This means that a connection between tone color and dynamics is almost non existent for these tones.

It is worthy of mention that played harmonics are especially easily addressed, and consequently can be played particularly softly: down to below a power level of 62 dB, which is softer than for violins and violas. The influence of mutes is also

stronger on the lower limit of the loudness region. A *pp* can be dampened down to 55 dB with a mute, without concurrently reducing the dynamic range: for *ff,* 89 dB are nevertheless possible, resulting in a variation possibility of 34 dB.

3.3.3.3 Time Structure

As is the case for higher pitched string instruments, the initial transients of higher frequency partials reach their final values more rapidly than the lower components. Furthermore, noise components arising from the attack of all resonances are especially pronounced at the beginning. Since the lower partials, however, exhibit significantly longer initial transients than in a violin, i.e., between 60 and 100 ms (Melka, 1970), it can happen for notes of short duration, that, in spite of a sharp staccato attack, only the high frequency partials reach their full amplitude, while the low partials radiate only weakly. As a result, as already suggested for violas, the tonal picture of fast passages lacks the desired sonority, and emphasizes the nasal and noise-like contributions excessively. In interaction with other instruments, this characteristic of the cello must be considered during fast staccato runs, so that wind players should not perform their staccato too pointedly. This should help to achieve a uniform expression for the passage. A typical example is given by the solo quartet of Flute, Oboe, Violin and Cello in the "Martyr-Aria" in Mozart's "The Abduction from the Seraglio," where both strings frequently have difficulty in adapting their tone production to the staccato of the flute and the oboe.

In contrast, for slow passages the cello tone can be attacked and developed very softly. In the low register, initial transients of over 300 ms occur in that setting. For the high A string the tone development takes about 200 ms.

The decay time for a cello is significantly longer than for a violin and a viola. This is caused primarily by the longer and heavier strings. When the bow does not remain in contact with the string at the end of the note, values of 2 s are obtained in the middle and lower registers and 1 s in the upper register for fingered notes. For open strings the decay time lies near 10 s for the C-string (since no resonance can extract energy from the fundamental) and near 5–8 s for the higher strings.

As is the case for the violin, for a *vibrato*, the frequency modulation caused by hand motion results in level fluctuations of individual partials. However, except for Wolf – tones, these fluctuations are not very pronounced in the low register, i.e., below 300 Hz. Furthermore, there are hardly any individual notable high frequency partials which are perceived as penetrating. The Performer's attention is therefore demanded primarily for partials in the range of 400–1,200 Hz when they fall on the edge of a very pronounced resonance.

3.3.3.4 Pizzicato

The large string length of the cello provides the performer with the possibility of varying the *pizzicato* over a particularly broad range of expression. It is thus possible to create a *pp* with a short duration sound power level of 51 dB. Such a

low level is rarely found for other orchestral instruments, particularly when realizing that in this context the strongest components occur at frequencies for which the ear is relatively insensitive. In contrast, for *ff* the same levels can be performed as *con arco*, for individual partials then lie in the region of 90 dB. A four string chord can even reach nearly 100 dB (recalculated after Burghauser and Spelda, 1971).

This large dynamic range naturally also involves great variation in decay time. In a *pp* individual partials can be heard for 50 to 200 ms, for high loudness a tone persists from 400 to 1,400 ms. For open strings these values move somewhere above 1 s, while for fingered strings they are shorter, depending on the length of the vibrating string portion (Spelda, 1968). If a very dry, i.e., short *pizzicato* is desired, this can be achieved by damping the string by means of harmonic fingering.

3.3.4   Double Bass

3.3.4.1   Sound Spectra

The double bass is among the lowest instruments of the orchestra. Four-string instruments reach down to E_1 (41 Hz), the lower limit of the tonal range for a five-string instrument lies near C_1 (33 Hz) or B_0 (31 Hz) (Planyavsky, 1984) depending on tuning. The air volume resonance, the lowest resonance of the instrument, lies, depending on instrument construction, between approximately 57 and 70 Hz, which is nearly an octave below the lowest note. Consequently the fundamental of C_1 has a level which lies about 15–20 dB below the strongest tonal contributions. This difference increases with increased air resonance frequency.

For the low double bass tones, the most important tonal contributions are located between 70 and 350 Hz. They provide the dark color and richness for the sound. However, they do not create a vowel-like character, since the formant region of the "u(oo)" begins above this range. A secondary formant near 500 Hz rounds out the tonal picture in the low registers with a tendency toward a dark "o(oh)." Even for *ff*, the spectrum above the bridge resonance, which lies near 1,250 Hz, drops with a slope of 15 dB/octave. In the upper pitch ranges the spectrum widens, where frequently an additional formant appears near 800 Hz, which in color lies on the transition between "å(aw)" and a dark "a(ah)." However, the frequency location of this additional maximum varies, depending on size and construction of the bass. As a whole, in the double bass, the tendency of a drop in power spectrum between the fundamental and the bridge resonance by 6 dB/octave, is evident as well, and furthermore, a resonance dependent waviness of about ±3 dB is superimposed on this uniform decrease.

Noise contributions to the tone of the bass frequently reach upper frequencies beyond harmonic partials. Thus a specific mixture appears in the tonal picture, often described as a "buzzing." This becomes especially prominent when the basses play alone, since this effect is mostly masked by simultaneous sounding of higher instruments.

3.3.4.2 Dynamics

When compared to a cello, the sound power level of the double bass as a whole is about 2 dB higher: The average *forte* lies near 92 dB, for *ff,* the double bass reaches a level of 96 dB for rapid tone sequences, and for individual notes it even reaches 100 dB, where especially the region around A_1 and A_2, as well as some notes in the third octave are supported by pronounced body resonances. At *pp* the levels move around 79 dB for fast note sequences, and 66 dB for individual notes, where the notes of the second octave can be played particularly sensitively. Thus, in practice, a dynamic range of 25–30 dB can frequently be realized. Especially critical are fast soft passages, for which the level of the double bass can hardly be lowered to *pp*, when all notes are to be addressed (Meyer, 1990).

The relationship between dynamics and tone color in the double bass is noteworthy. Only in the upper registers does the overtone content rise as the dynamic level increases: As is the case for other string instruments, the "dynamic tone color factor" has a value of 1.1. However, in the first octave it drops to 0.8, and in the second octave to as low as 0.6. This means that the strongest tone contributions change more than the higher components. The reason for this lies in the fact that for low playing volume the bow pressure is reduced so much, that the fundamental vibration of the string is no longer fully developed. The tone thus loses substance and becomes more tender and gentle, and can also obtain a nasal timbre.

The greater loudness of the double bass in comparison to other string instruments also appears for harmonic fingering, for which the power level can rise to 91 dB at *ff,* and can be weakened to 74 dB for *pp*. However, the overall level of the double bass can be reduced especially by employing a mute. The reason for this appears to be in the fact that the tone contributions which are effectively damped by the mute, lie in the frequency region of large ear sensitivity, while the low tone components in the region of relatively closely spaced equal loudness curves (see Fig. 1.1) decrease noticeably in their perceptive impression even for a smaller decrease in level. The dynamics of a muted double bass cover a region of about 68–88 dB (recalculated after Burghauser and Spelda, 1971).

3.3.4.3 Time Structure

As a result of its size, the double bass, in its low registers, requires significantly more time for the initial transients than the higher string instruments. Even for sharp attacks, the initial transients in the region of C_2 require over 120 ms, only from C_3 on up are values of below 100 ms achieved (Melka, 1970). However, the higher tone contributions have a shorter initial transient, so that in an even stronger measure than in the cello, a tone color change results for notes of very short duration: The low components, which otherwise determine the fullness of the tone, remain too weak, and the contributions in the region of nasal formants determine the tonal impression. Furthermore, for fast staccato passages, the relatively strongly

pronounced noise development during the attack becomes especially noticeable, yet in a positive fashion, this brings about a rhythmic articulation of such phrases.

For a broad tone development, the initial transient in the upper register of the double bass requires about the same time as in other string instruments (150–250 ms), in the lower registers the spectral development requires more time. Thus the notes around C_3 can reach initial transient times of about 350 ms, and near C_2 even more than 400 ms. Open strings, however, do not permit such a soft attack, even the C_2 string does not have initial transients longer than 180 ms (Luce and Clark, 1965; Melka, 1970).

The decay times of the double bass are slightly longer than for a cello. For fingered notes they are of the order of 3 s, and for open strings near 10 s. As noted in Fig. 2.10, these data refer to the lower partials. The decay time of higher partials lies in the region of 0.5 s.

3.3.4.4 Pizzicato

In spite of the low frequencies, the pizzicato tones of the double bass have extraordinarily short initial transients. The open C – string requires only about 35 ms. For higher positions the initial transient time drops steadily: for C_2 it is still about 25 ms and drops to less than 15 ms above C_3. Thus in all regions, a very pronounced tone placement is possible (Melka, 1970).

Dynamic limits lie about 3 dB higher than for the cello. A sound power level of 93 dB can be achieved for short durations at *ff,* for *pp* this can be reduced to 60 dB, which corresponds to a range of 33 dB; accordingly, the dynamic possibilities for a *pizzicato* are greater than for *con arco* (Spelda, 1968).

With a maximal perceived tone duration of 1.6 s at *ff,* the double bass exceeds all other string instruments. However, this value applies only for the three middle open strings, yet a median value of at least 1 s can be assumed for all registers. For notes, which are fingered in the higher registers, the decay time is reduced by about 2/3 of the value for the same note on the next higher string. Even when playing very softly the notes continue to sound for 400–500 ms, which is significantly longer than for a cello.

As is the case for the other string instruments, the low tonal contributions have a significantly longer decay time than the higher ones. This results in a very homogeneous tonal effect in *pizzicato* for the entire string section, since this decrease in decay time in relation to the spectrum finds a parallel in the decay time decrease in the higher instruments. For *pizzicato* – chords which include all string groups, a uniform tendency of frequency dependence in decay time results, which can lead to a bell-like effect. A pronounced example for such a tonal effect is found in the 5th Symphony of P. Tchaikovsky, an excerpt of which is represented in score example 9, where, in spite of the dynamic indication of *mf*, the skilled compositorial use of open strings enables a wide vibrational amplitude for each attack. Yet it is all the more astonishing to experience orchestras which deliberately avoid open strings by appropriate fingerings and *divisi*-performance, in order to create as soft a tone as possible, which likely does not correspond to the intent of the composer.

Score example 9 P. I. Tchaikovsky, Symphony No. 5, 2nd Movement, measure 108 ff

3.4   The Piano

3.4.1   Sound Spectra

No stationary state is created for the piano sound, since there is no uniform continuous excitation. Nevertheless, quasi-stationary conditions can be assumed as an approximation at least for short durations. As a result, spectra of partials can certainly be used for the tonal description of the sound during its initial phases, however the time structure, and above all the decay behavior play a much more important role than in string and wind instruments.

The sound spectra of the piano are in large measure determined by the sound radiation characteristics of the sound board, which has its strongest resonances in the region of about 200–1,000 Hz. For larger instruments, resonances can even appear between 100 and 200 Hz, which lend additional fullness to the lower register of the tonal range. Furthermore, depending on instrumental construction details, resonance effects between 100 and 200 Hz can be intensified, which aids the brilliance of the tone (Wogram, 1984).

Accordingly, for the largest portion of the piano sounds, the fundamentals of the partial spectra dominate. Only in the two lowest octaves of the tonal range, the intensity maximum is shifted to overtones in the frequency range of about 100–250 Hz. Below this amplitude maximum, the strength of the partials decreases with a slope of from 12 to 15 dB/Octave, so that the fundamental of the lowest note A (27.5 Hz), with a level of 25 dB, lies below the strongest component.

Above this amplitude maximum, i.e., in the middle and upper registers, the envelope for most pianos decreases quite steadily. The average level decrease of the envelope below 1,500 Hz is of the order of 10 dB/Octave, individual partials in contrast, depending on their location near resonance peaks, or between resonances, can be enhanced or attenuated by several dB. Above 1,500 Hz, the spectra drop by 15–20 dB/Octave. For tones near C_8 (4,100 Hz) this means that the spectra only contain three partials, where the octave component is so weak that the sound looks almost like a pure sine wave.

Tone color determining formants are rarely found in grand pianos or uprights. They only appear in very few models between 500 and 2,000 Hz. In contrast to grand pianos, upright pianos have a tendency to emphasize certain frequency ranges between 100 and 350 Hz. This is based on resonance effects of air enclosed in the piano housing and, depending on the location of the piano in the room, similar resonance effects of air between the wall and the piano sound board. These lend coloration to the tone which is independent of the fundamental tone. It is particularly noticeable in multi voice playing. A grand piano on the other hand is more suitably adapted to variations in register (Bork, 1992).

A peculiarity of the spectra for some models, or at least for some limited tone range, consists in the fact that certain partials are suppressed, or are reduced in level by locating the hammer impact point at a whole number fraction of the string length. For example it is often attempted to weaken the seventh partial and its multiples by locating the hammer impact point at 1/7 of the string length. The object is to avoid the roughness of the seventh. This concept, however, is only fully effective for short tones, since the amplitudes of these partials adjust themselves to neighboring partials by mutual energy exchange for vibrations of longer duration. (Meyer and Lottermoser, 1961).

Two further characteristics, which participate in the development of the piano sound timbre, are the inharmonicity of the overtones, and the noise contributions. The latter are significant primarily during the attack, however, they are also observable in the spectra. Their frequencies reflect the resonance distribution of the instrument. Thus, the strongest noise contributions are to be expected in the range from about 300 to 750 Hz. In the highest registers of the tonal range, their level can reach up to 6 dB of the fundamental of the partial spectrum (Wogram, 1984). Inasmuch as the frequency of the noise lies far below the fundamental, it is no longer masked by the actual tone, and thus becomes particularly noticeable.

In contrast to instruments with stationary excitation, the overtones of the piano sound do not have a strict harmonic frequency location, but are stretched, i.e., they are located slightly higher than the whole number multiples of the fundamental frequency. This stretch is especially evident in the upper registers, however, it can also be noticed in the mid and lower registers. The amount of the stretch does not differ significantly between grand and upright pianos, so that inharmonicity, contrary to earlier opinions, does not make a contribution worthy of mention, to the general tonal difference between them (Bork, 1992). Rather, it represents a tonal characteristic for both instrument types.

Finally, longitudinal vibrations need to be mentioned. They can lend a unique coloration to the lowest registers. Usually the longitudinal partial, which is most frequently inharmonic, lies between the 12th and 20th partial of the harmonic spectrum. Its frequency location cannot be influenced by the usual tuning procedures. It is thus located in a region of good sound board amplification as well as high ear sensitivity. As a result it is perceived as disturbing when playing scales, since the longitudinal partials for different strings occur at different frequencies and are thus perceived as nonsystematic frequency jumps (Bork, 1989; Conklin, 1990).

Relief, i.e., shifting longitudinal partials to harmonic frequencies, is only possible through appropriate choices of string materials.

3.4.2 Dynamics

While the dynamic range of the grand piano is primarily determined by the key attack, it is also influenced by the use of the pedals and the position of the lid. When playing scales in two voices the sound power level at *ff* can reach around 104 dB. In the low registers it can be 1–2 dB higher than in the upper registers. The values apply with open lid, without the use of the right pedal. Use of the pedal raises the power level in the low register by 4 dB, in the upper register by 3 dB. Closing the lid, on the other hand, lowers the level by only 1–2 dB. In *pp* dual scales lead to a power level of the order of 88 dB without essential influence by register. Use of the left pedal lowers the level only by 1 dB, closing the lid effects a further decrease by 2 dB. For individual notes the *pp* can be further reduced, so that it is possible to drop to a sound power level below 65 dB. When one further considers the level increase due to a full two handed performance, the dynamic range of a grand piano rises to roughly 45 dB.

The felt surface of the hammer is increasingly hardened by contact with increasing strength of attack. Consequently, the exciting impulse becomes richer in overtone content with rising dynamics (Askenfelt and Jansson, 1990). This means, that for the spectrum, an increase in 1 dB of the strongest partials, a corresponding rise of about 2 dB (for some grand pianos as much as 2.5 dB) is observed for overtones in the region of 3,000 Hz, so that the tone of the open grand piano gains both in brilliance and brightness. When closing the lid, however, the higher frequency contributions are damped about twice as much as the strongest partials.

3.4.3 Time Structure

Inasmuch as the excitation of the strings is caused by the impact of the hammer, the speed of impact, on the one hand, and the contact duration, on the other, determine the tonal development, where the hardness of the hammer felt also plays a role. For upright, and grand pianos, therefore, an initial transient results, the duration of which has the same order of magnitude as is found for a pizzicato in bowed string instruments. In the lower registers the initial transient lasts around 20–30 ms; for individual notes, favored by resonance conditions, this is reduced to 15 ms. In the upper register the initial transient is reduced to values between 10 and 15 ms (Melka, 1970).

The attack noise is a characteristic peculiar to the piano. The example of the initial transient for the C_6 in Fig. 3.21 shows various noise components in addition to the partials. It is noted that the low resonances of the instrument are excited, they vibrate over a time period of 100 ms, or even more, and give a certain color to the tone. Furthermore, between 600 and 2,500 Hz, a clicking noise of short duration is
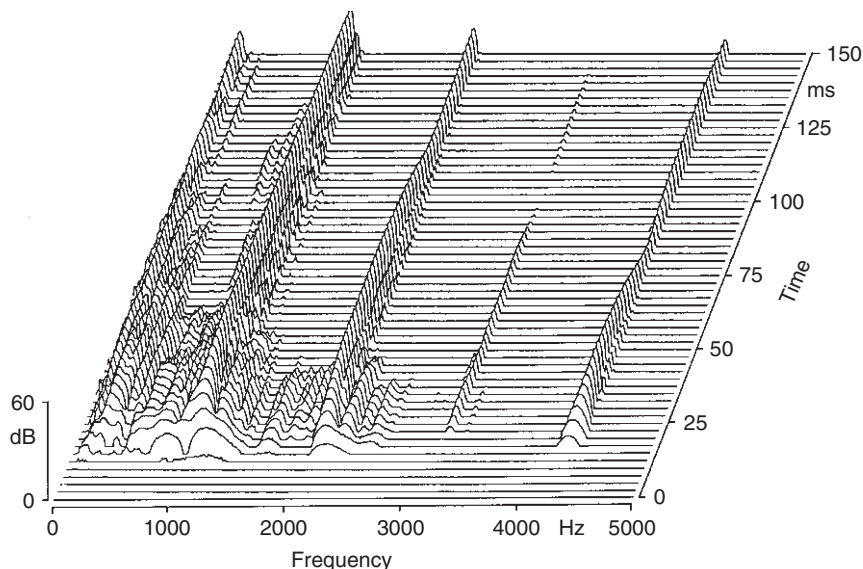
Fig. 3.21 Time evolution of a tonal spectrum for a grand piano (played note C_6)

recognized. It only lasts for 25–40 ms, and provides the attack with a certain kind of articulation. Finally, the actual partials are accompanied by side bands, which are caused by a very brief excitation of neighboring strings. These, however, are masked in the perception processes in the ear.

Noise components in the attack provide important contributions to the tonal variation possibilities obtainable by the nature of the attack. They result from the various motion processes in the keyboard mechanism. While, during a legato attack, the key is accelerated uniformly, the staccato motion of the key is characterized by layered fluctuations, which are transmitted on to the hammer motion (Asklenfelt and Jansson, 1990). However, the high speed of the key motion is also transmitted to the frame and sound board in the form of a force impulse through the key support, so that even before the hammer impact on the string, an audible noise can be created (Askenfelt, 1993). Figure 3.22 shows the initial transient for a staccato and a legato attack (Koornhof and van der Walt, 1993), the time of hammer contact with the string is given by "0." This representation, where a coarse frequency resolution was chosen in favor of better temporal resolution, clearly permits observation of the preliminary knocking noise, the duration of which is barely 40 ms. Listening tests in a concert hall have shown that the presence or absence of such articulation noise is clearly discernable. The assumption for this is of course, that this noise is not masked by an earlier tone. In this sense, legato performance, rests in large measure on the temporal overlap of tones by the nature of the connection, including pedal technique.
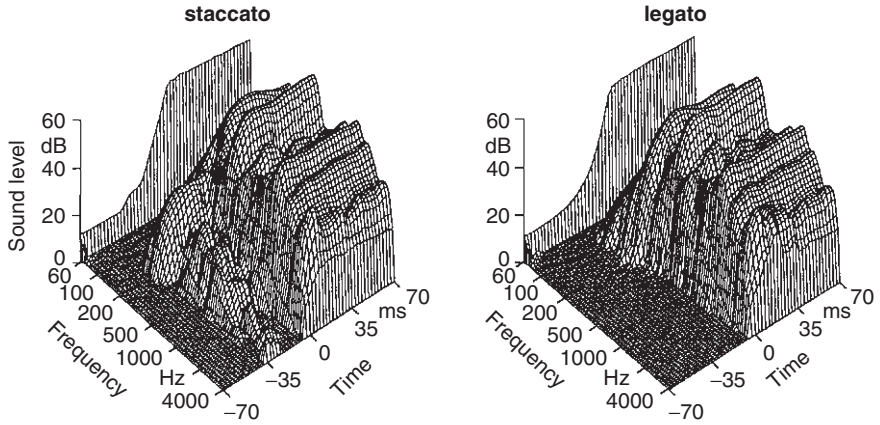
Fig. 3.22 Influence of attack technique on the noise components of initial transients for a grand piano (played tone: C_3; after Koornhof and van der Walt, 1993)
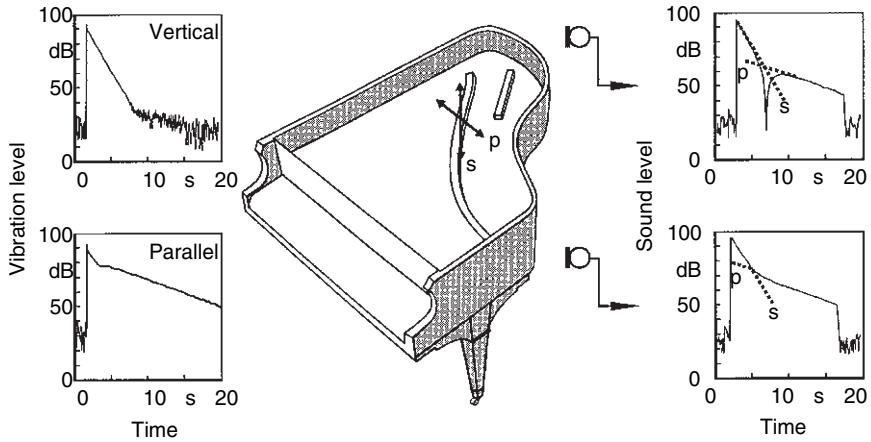

Fig. 3.23 Decay behavior of a grand piano (after Weinreich, 1977). *Left*: Sound board vibrations in two planes, *Right*: radiated sound at two microphone positions

The temporal fine structure during the decay belongs to the most important tonal characteristics of the piano sound. When the dampers are lifted from the strings, i.e., by depressing the right pedal, the sound can be followed for 10 s or more. Initially the intensity decreases more rapidly, and subsequently decays over a longer period of time in the region of decreased loudness. The reason for this can be explained in the context of Fig. 3.23: the initial impact on the string occurs in a direction perpendicular to the sound board; in this direction the sound board is in a position to extract energy from the string in relatively strong measure, as shown in the upper left partial picture indicating the amplitude of the sound board vibrations. In addition, string vibrations parallel to the sound board are formed, though much
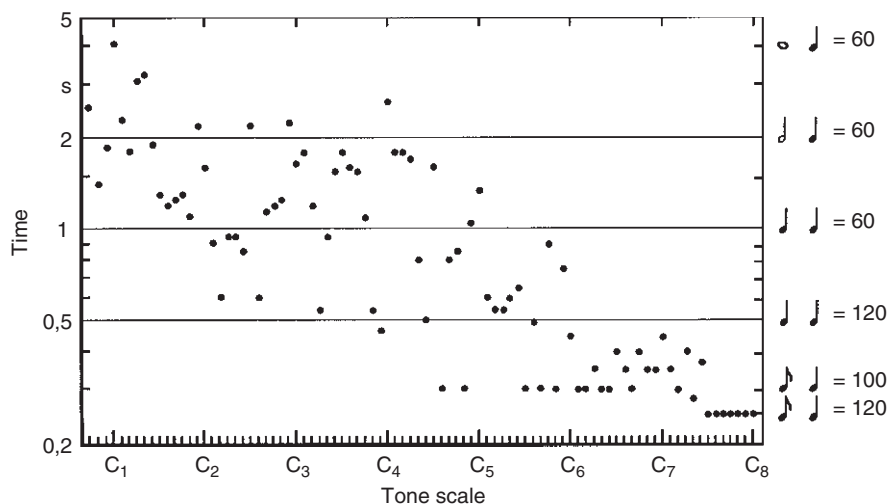
Fig. 3.24 Duration of the first, linear decay phase for individual notes of a grand piano (after Meyer and Melka, 1983)

weaker. Since the sound board presents a much higher impedance for transmitting such vibrations, this energy transmission process is much slower (left lower partial image). The radiated sound field includes a superposition of these two different forms of vibration. Depending on the relative phase of the two components the decay process of the sound field surrounding the instrument can produce a different time structure (Weinreich, 1977). In principle the time evolution of the decay process can be represented by two straight lines. The first of these drops with a relatively steep slope from the maximum amplitude, while the second continues the decay with a noticeably shallower slope. This time structure is independent of the dynamic level of performance.

The decay of short notes is determined exclusively by the slope of the first decay phase. The duration of this first linear phase is represented in Fig. 3.24 for all notes of a grand piano, where the time scale is indicated by note values and metronome indications at the right side of the graph. The strong scatter of individual points is noteworthy. An explanation for this is found in the variations of the initial slope, and also in the variations in the level for the onset of the subsequent slower decay. This contributes significantly to the tonal animation of the instrument. In general it can be said that in the low register the first decay phase corresponds to a half note or at least a quarter note at a slow tempo, while in the mid register it is sufficient for fast quarter notes, and is limited in large measure to allegro eighths notes in the upper register.

The determination of the decay time for this first decay phase is best based on the slope for a level decrease of the first 10 dB (subsequently recalculated for 60 dB). This is in analogy to the "Early-Decay-Time" in room acoustics (see Sect. 5.3). Since this initial decay time can differ by up to a factor of 2 for neighboring notes,
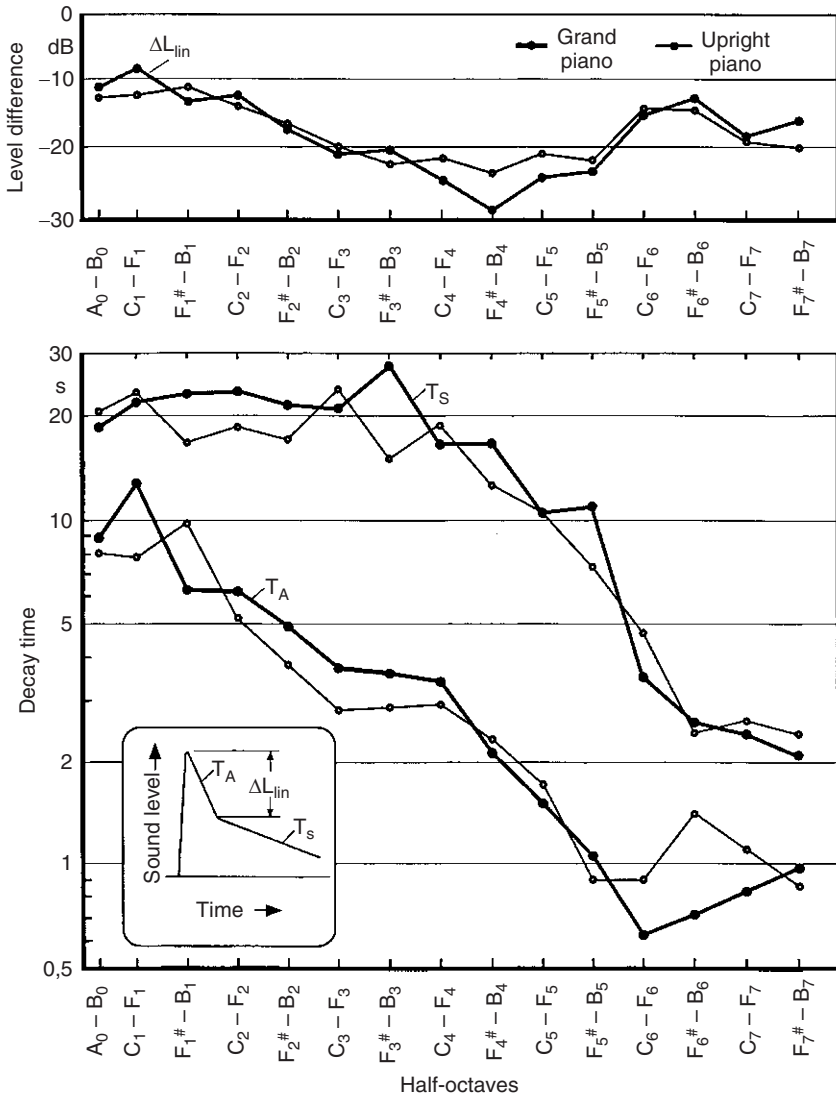
Fig. 3.25 Decay times for grand pianos and upright pianos (Median values for half octave regions). *Top*: Level difference between peak level and onset of the later decay phase. *Bottom*: Initial and later decay time

the typical sequence for a tone scale is represented in Fig. 3.25 in such a manner, that in each case a median value was formed for all notes within half an octave: Starting with values of 10 s, the initial decay time drops with rising pitch by a factor of 1.7 per octave. This results in values around 3 s in the middle register, and between 0.6 and 1.4 s in the highest registers. Grand pianos and uprights exhibit the same tendency (Meyer and Melka, 1983).

The duration of the initial decay time is primarily responsible for the song-like characteristic of the melody line and the connection between the notes of flowing music. When individual notes of an instrument exhibit particularly short decay times, they give a dull and dry impression and they drop out of the overall tonal picture. The threshold for detection of differing initial decay times by the ear depends on their duration. From a simplified approach it can be said that for decay times in excess of 4 s, changes of about 25% are necessary to detect a difference. For decay times of between 1.5 and 3 s, changes of 15% are sufficient, for decay times of less than 1 s, changes of 10% already suffice.

Notes of long duration are needed for the subsequent temporal fine structure of the decay process to gain significance. The essential characteristics of the later decay time (again referred to a level decrease of 60 dB) and the level difference between the peak value and the break value are represented in Fig. 3.25. On the average, the values of the later decay time move around 20 s for the lower half of the tonal range, where individual notes can have a decay time of as long as 30 s. In the upper register the later decay time is reduced by a factor of about 1.9 per octave, and thus reaches a value of between 2 and 3 s for the highest notes. Individual differences between grand pianos of different manufacture result from the fact that the longest values of decay time occur in different registers. Thus for instruments which emphasize the low notes the longest decay times fall into the 2nd octave, while others, by preference of the mid register, have a more rounded, but in the whole, less-full sound.

The level difference between the peak value and the breaking point separating early and late decay is shown in the upper portion of Fig. 3.25 as additional decay information. Again, mean values of half octaves are considered. The dB values of adjacent isolated notes can vary by up to 30%. The threshold of hearing distinction lies at around 3 dB. It is a characteristic of almost all instruments that the slower decay in the low registers begins at a relatively high level. The breaking point, however, drops steadily up to the mid register, so that the first decay phase clearly shows a linear drop of over 30 dB within the middle (C_4) octave. The tone thus gains clarity, without causing the flowing character to suffer, assuming a sufficiently long decay time. In the upper registers the level of the slower decay rises noticeably. Because of the shortness of the decay time, this does not detract from the clarity of the tone picture, and furthermore, it supports the tonal content of the higher notes. In addition, particular emphasis needs to be given to the fact that the level difference between the maximum and the transition from the more rapid to the slower decay does not depend on the strength of the attack. In contrast, the level of the slower decay can be raised by several dB by activating the left pedal. This enhances the effect of the decay.

The representation of the decay process by use of a broken line (as in the sketch of Fig. 3.25) does not yet give a complete description of the time dependence of fine structure. It is true that level trends of that kind are the norm in low registers, yet in the upper registers, level curves frequently show a superposition of beat-like variations onto the regular level decay curve. Already in the mid-registers from 30 to 50% of the notes are affected by this. From the C_5 octave upward this includes from 70 to
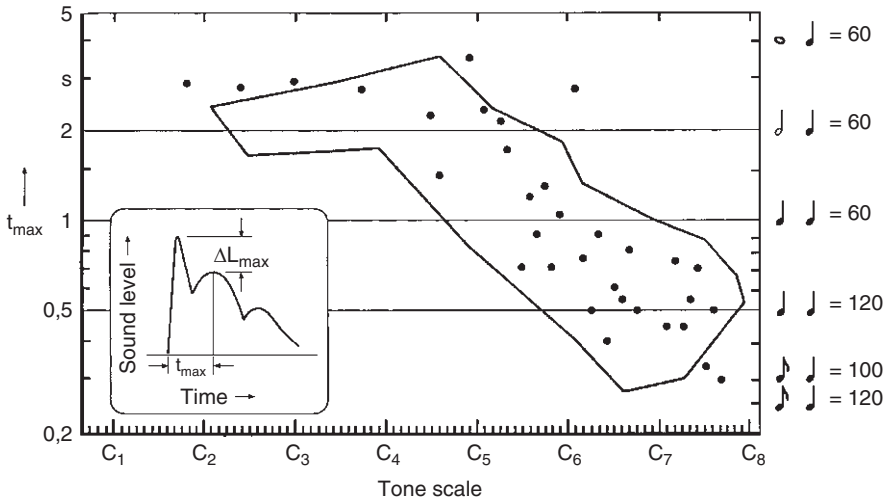
Fig. 3.26 Time location of the first beat maximum in the decay processes for a grand piano (measurement points for individual partials) and an upright piano (enclosed region of spread)

over 90%. The first maximum of these "beats" follows the tone onset after about 2–3 s in the lower half of the keyboard, for higher registers this delay is reduced by a factor of about 2 per octave. As shown in Fig. 3.26, this breathing of the tone is only noticeable in the lower registers for relatively long notes, while in the upper register it already becomes significant for rather rapid tone sequences. In general, this first "beat" maximum lies about 10–20 dB below the level of the initial peak, however, for individual notes this level difference can assume values between 8 and 35 dB (Meyer and Melka, 1983). It should also be noted that the temporal level fluctuations occur even for well tuned pianos. This is caused by changes in the direction of string vibration, furthermore, differences in inharmonicity between the three strings associated with the same note can play a role.

3.5 The Harpsichord

3.5.1 Sound Spectra

In the harpsichord, as in the piano, the frequency range of the most strongly radiated sound intensities are determined by the main resonances of the sound board. They lie in the range of 200–800 Hz. Depending on construction details, the maximum of the radiated sound is found between 300 and 600 Hz. As determined by the process of string excitation through plucking, the string excitation is much richer in overtones than for a piano. In addition, the inharmonicity of the strings is significantly less than for piano strings. In the low registers of the harpsichord, at the 50th

partial, a deviation from the harmonic frequency of about 15 cents is usual, while a deviation of 30 cents is already considered to be relatively high. For the same pitch on the piano already the 6th partial lies 30 cents above the harmonic frequency.

Below 200 Hz, the radiated sound energy drops with a slope of more than 40 dB/octave, so that only very weak partials can be expected. Consequently in the lower registers the strongest partial can be found between 200 and 500 Hz, while the fundamental always dominates in the mid- and upper registers. The intensity distribution at higher frequencies is in large measure time dependent. For a numerical description it is therefore recommended to base energy content on the first second of the harpsichord sound (Elfrath, 1992). Above 800 Hz the spectra drop initially with a slope of about 7 dB/octave, and subsequently pass through a dip between 2,000 and 2,500 Hz, which lies barely 15 dB below the strongest partials. This is followed by a secondary maximum, which lends its particular presence to the harpsichord. Above about 5,000 Hz the spectrum drops at a rate of about 15 dB/octave.

In comparison to the otherwise minor spectral differences, this formant, which lends a certain presence to the instrument, is rather individually pronounced for harpsichords of different construction. The rise, in contrast to the previous dip can vary between 2 and 6 dB, so that finally the secondary formant lies about 7–12 dB below the strongest spectral components. This frequency varies also. It is most often located near 4,000 Hz; for some instruments, however, around 5,000 Hz. However, it is always situated clearly above the so-called singer's formant (see Sect. 3.8.1.), which, in its significance for the presence, and the ability to carry the sound of the voice, has a similar function. As an example, this effect is clearly observable for a harpsichord continuo within an orchestra, when these high frequency contributions, in their rhythmic structure, stand out from the overall sound, while the harmonic chord foundations are only of subordinate importance or are even totally inaudible – at least to the listener in the hall.

3.5.2   Dynamics

The dynamic range of the harpsichord is very limited, since the nature of the key attack has no essential influence on the process of plucking the string. Different dynamic steps are therefore only accessible through registration, i.e., by playing one or more strings for each key. The combination of two 8′ registers results in a power level increase of 2–3 dB in comparison to the single register, combination with a 4′ register effects a change in tone color in the direction of a brighter timbre, which also gives the impression of a dynamic increase. In the mean, one can count on a power level (calculated after Burghauser and Spelda, 1971) of between 71 and 87 dB, depending on registration and performance technique. These are values which cannot only be reached by a single violin, but can certainly be exceeded. In comparison it should be mentioned that the clavichord is even softer by about 10 dB.

The low power level of the harpsichord frequently leads to the desire for electroacoustic amplification, to improve the dynamic balance with the remaining ensemble. It is important for such a reproduction of the harpsichord sound with

speakers, to counteract a nasal sound caused by an additional lowering in the frequency region around 1,500 Hz, when the harpsichord is amplified above its original loudness (Thienhaus, 1954).

3.5.3 Time Structure

The duration of the initial transient of individual harpsichord tones is very short. In the middle and upper registers it is only 10–25 ms. In the low registers this can be stretched to the range of 45–75 ms, depending on structural characteristics of the instrument (Neupert, 1971; Weyer, 1976). In addition, an articulation peak of short duration, approximately 20–30 ms, occurs with a principal intensity mostly above 2,000 Hz. This lends a degree of precision to individual harpsichord tones which can lead to an uncomfortable hardness in the performance of several simultaneous notes. This is at least one of the reasons why chords on the harpsichord are usually performed somewhat arpeggiated.

The superposition of vertical string vibrations and vibrations parallel to the sound board determines a decay process similar to that in the piano. However, the initially steeper amplitude drop is less dominant for the overall tone than it is in the piano, since on the one hand the parallel vibration is more strongly excited by the plucking process, and on the other hand the vibrational energy is transferred more rapidly from the vertical to the parallel vibrations. The temporal division of the decay process is thus not necessary (unlike the piano) and the specification of a single value for the decay time is sufficient.

In Fig. 3.27, the decay times of four historic harpsichords of different styles are represented in relation to tone location (Elfrath, 1992). On the whole, values are found, which in order of magnitude can be compared to the late decay times of
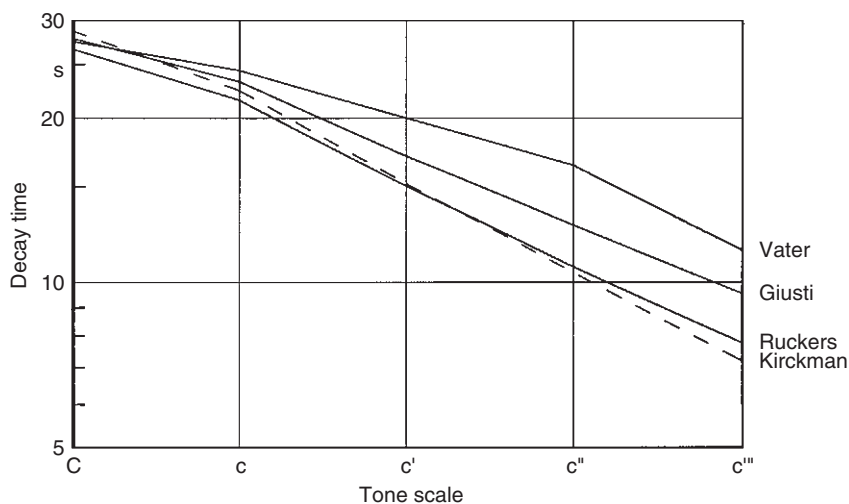

Fig. 3.27 Decay times for four harpsichords of different origin (after Elfrath, 1992). A. Ruckers, Flanders 1628; G.B. Giusti, Italy 1681; Ch. Vater, Germany 1738; J. Kirkman, England 1783

pianos; not, however, to piano initial transients. In the low octaves, the curves for the individual instruments are located very close to each other, but drop subsequently with differing slopes (by a factor of 1.25–1.5 per octave), so that the decay times differ by up to 70% in the highest register. Even though these curves represent mean values, where, for each instrument, decay times can be clearly differentiated from tone to tone, they show the individual characteristic of the instruments more clearly than the time averaged spectral composition. In addition, the different decay times of the higher overtones lead to a change in tone color within less than 100 ms. This also contributes to the characterization of the instrument.

Newer harpsichords have decay times closer to the Kirckman instrument in Fig. 3.27, where the possible spread is relatively large. To generalize, one can stipulate a decay time of about 20 s for C_3 with a drop by a factor of 1.4 per octave (see also Neupert, 1971; Fletcher, 1977). Assuming an initial sound power level of 70 dB and a noise level in the room of 30 dB, the tone can be followed by the ear for about 2/3 of the decay time. The real tone duration, however, is almost always significantly shorter than the decay time, so that the tonal character is essentially determined by the level drop of the first 10 dB or at most 20 dB.

3.6 The Harp

3.6.1 Sound Spectra

Sound radiation by the harp is essentially determined by a few distinct body resonances. The three most important resonances lie between 200 and 450 Hz. These are followed by two additional strong resonances up to 850 Hz (Firth, 1977). Above 1,000 Hz, the resonance sequence is more closely spaced, however, the individual resonances steadily decrease in strength. Inasmuch as the tonal range of the harp extends down to a C_1^b (31 Hz) the fundamentals of the spectra are relatively weak in the low register, the overtones which fall on the resonances between 200 and 450 Hz dominate. In the range of from G_3 or C_4 upward the fundamental becomes the strongest component of the spectrum.

As is the case for *pizzicato* playing on string instruments, the overtone content of the harp sound depends on the plucking location. For a pluck near the center, a complete spectrum is formed with strongly decreasing overtones, thus in the mid and upper registers, already the octave partial is from 10 to 15 dB below the fundamental, while in the low register, the decrease of the spectrum begins above 450 Hz. For an attack exactly in the middle of the string, the odd partials clearly dominate in contrast to the even ones, and the sound becomes full and soft. An attack at 1/3 of the string length suppresses the third partial and lends brightness and brilliance to the tone through the relatively strong octave partials (2nd and 4th order). Attack near the end of the string *(presso la tavola)* leads to a spectrum which drops only by 20 dB for the first eight partials, and thus has a guitar-like or possibly even metallic sound.

3.6.2 Dynamics

The harp and the piano are similar, in that the sound power radiated by the harp can only be indicated for the initial time period of the tone, i.e., in the time region of the strongest sound development. At the lower limit of the dynamic range the sound power level lies at around 60 dB (as calculated after Burghauser and Spelda, 1971), only when reaching the second octave does it rise to a level of about 70 dB. The upper limit rises from about 88 dB for the lowest notes to 100 dB in the region of C_4 and then again drops to about 80 dB in the C_6 octave. This results in a maximum dynamic range of 40 dB around C_4.

3.6.3 Time Structure

The attack for the harp is characterized in large measure by a sharp precision. This is partly determined by very short initial transients. For the lowest notes they are about 20 ms. In the upper register they drop to less than 10 ms (Melka, 1970). Tones, for which the fundamental falls on a resonance, have an initial transient of longer duration than neighboring tones. The coincidence of the neighborhood of a fundamental with one of the five main resonances can lead to short term beats during the initial transient.

The second reason for the precision in the initial transient can be found in the decay behavior of the harp. The higher partials, caused by the precise attack decay much more rapidly than the lower partials resulting from the full tone. In the lower register these have a decay time of the order of 4–6 s, in the mid-register about 2 s. Tones, whose fundamental falls directly on one of the main resonances have a noticeably shorter decay time than their neighbors, which causes the relevant tones to become dull and blunt.

Because of usual string lengths and strengths, a string in its fundamental tuning is longer and thinner than its neighboring, next lower string, raised to the same note by a pedal shift. This enables a tonal differentiation. Strings in fundamental tuning decay more gently, strings with raised tuning result in a harder sound (*"secco"*). The particular performance technique of increasing the decay time rests on the ability to tune two strings to the same pitch by a pedal shift, to couple them acoustically. For example Puccini specified this effect in "Turandot."

Inasmuch as "unused" strings are not damped, in contrast to the piano, they experience mutual coupling to the vibration processes through the resonant body. In this function they contribute significantly to the increase in decay time. In a given setting they must be damped by hand to interrupt the decay. It is also possible to excite string vibrations by a strong external sound field through the sound board. Orchestral chords can initiate such a decay without harp participation. This is clearly audible, at least in the neighborhood of the harp, and possibly needs to be damped. This is particularly important if there is a microphone located in the vicinity of the harp.

3.7   Percussion Instruments

3.7.1   Timpani

For percussion instruments, even more so than for the piano, the tonal character is largely determined by the time structure of the spectral composition. On the one hand it could be of interest to consider the tonal spectra at the instant of strongest sound radiation, on the other hand, the different decay of individual spectral components plays an important role. This is particularly clear for tympani, where after initial impact noise, harmonic components dominate, evoking a clear pitch impression.

The membranes of timpani can vibrate in a multiplicity of different vibrational shapes (so called vibrational modes). For the simplest mode, the rim forms a nodal line, and the membrane vibrates with its entire surface area in phase. Added to this lowest "ring mode" are higher modes, for which additional nodal lines are formed as concentric circles. A further group of modes is formed by nodal lines running radially across the membrane which cross in the middle ("radial modes"). This group has the important property that the frequencies of the first three to five modes are in reasonably good harmonic relation to each other: their frequencies ratios approach the numerical ratios of 2:3:4:5:6; these modes thus are essentially responsible for the pitch impression (Rossing, 1982b; Fleischer, 1991).

The perceived pitch lies an octave below the main tone, that is for the large concert kettle with a tonal range of F_1 to D_2 between about 44 and 73 Hz, and for the small concert kettle (A_1 to G_2) between 55 and 98 Hz. The low D kettle reaches down to D_1 (37 Hz), the high A kettle up to C_3 (130 Hz). However, pitch perception is not as precise as it is for string and wind instruments, because partials are not as accurately harmonic, and the frequency location is so low. Consequently, Verdi for example, maintains the original tuning for complicated modulations, when time does not permit retuning and uses the notation of G for the tonic in G^b major. This problem has been eliminated in the meantime by the invention of the pedal tympani.

Timpani tuning is accomplished by changing tension in the membrane. For a good instrument, this changes the frequencies of the radial modes in proportion to each other, i.e., harmonic relationships are retained. This property of timpani is related to the kettle size. Its volume influences the vibrational frequencies of the membrane within certain limits. The lowest ring mode experiences a smaller frequency change than the radial modes, thus it does not move harmonically with the other modes when tuning.

The strength of the partials radiated by these vibrational modes depends on the location of the impact. Impact near a nodal line is almost totally ineffective in exciting the corresponding mode. Impact at the center of the membrane would therefore excite the (inharmonic) ringmodes strongly, and the (harmonic) radial modes only weakly. In contrast, the usual impact location, located more near the rim favors the harmonic components and reduces the inharmonic ones. The player can selectively excite the lowest radial mode (the fundamental) or the next higher radial mode (the fifth) most strongly.
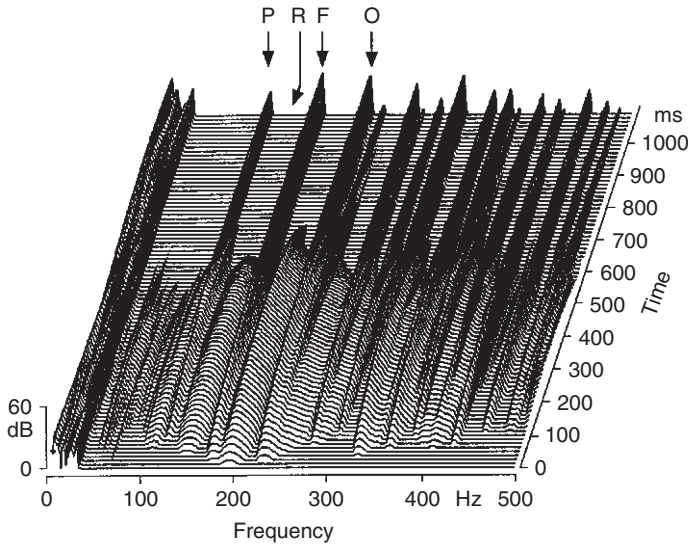
Fig. 3.28 Time evolution of a timpani spectrum of pitch A (after Fleischer, 1991). P Principal mode, R 1st Ring-mode, F Fifth, O octave

Figure 3.28 shows the change of the spectral composition of a timpani tone with time. Approximately the first second of the sound is represented. For the duration of approximately the first half second, many quickly decaying components are recognized. These constitute the impact noise. In this example the lowest ring-mode lies at around 140 Hz. After that, a series of slowly decaying partials remains. The principal mode at 110 Hz, as well as the fifth and the octave stand out prominently. Usually the decay time of the principal mode moves between 7 s for the low and 1.5 s for the highest register. The decay times of 10 s and 3 s for the fifth and the octave respectively are significantly longer, so that these partials, in time, dominate over the principal mode. For natural skin membranes this difference is not as pronounced as for man made materials. For natural membranes the pitch dependence of decay times is also less pronounced, consequently, as a whole, it gives a more even impression (Fleischer, 1991). With hand damping the decay time is reduced down to 0.7 s for low, and to 0.2 s for high pitches.

The strong harmonic partials mask the inharmonic components relatively early in the auditory impression, contributing to the purity of the timpani sound. This is the case, provided that, – aside from the correct impact location, the frequency of the lowest ring mode does not lie below the fundamental. This condition is no longer satisfied for high tuning of the drum head, therefore the high timpani tones sacrifice clarity (Fleischer, 1991). The timpani sound can also lose clarity, when the vibrational modes become less precise due to uneven membrane thickness. In that case the tonal character depends especially strongly on the point of impact, a circumstance which the player can utilize for particular nuances.

The initial transient is characterized by the relatively slow development of the low pitch contributions, which move in the neighborhood of 100 ms. At the same time, higher frequency contributions can prefer the entrance point at the time of the timpani impact to some extent, if they are sufficiently pronounced. For this, Melka (1970) indicates initial transients of less than 20 ms.

The playable dynamic range of about 45 dB is very large. For ff a sound power level of 115 dB can occur. For *pp* this reduces to 67 dB in the low registers and down to 70 dB in the upper registers (recalculated after Burghauser and Spelda, 1971). The strongest tonal contributions develop between about 100 and 250 Hz, depending on tone location and impact characteristics.

3.7.2   The Bass Drum

In contrast to timpani, the bass drum belongs to the unpitched percussion instruments. The muffled tone, with its strongest contributions in the frequency range near 100 Hz, is characterized by a multitude of inharmonic partials, some of which are closely paired. The usual soft mallets prevent excitation of higher frequency components. The heavier the mallet, the more energy can be transferred to the membrane, however, this also increases the contact duration, which in turn suppresses the higher frequency components. In addition, the more closely centered impact location leads to a preference of the ring-modes of the membrane vibrations (if possible, these modes are to be avoided in the timpani), which likewise supports the dull non-distinct tone character. For an impact point closer to the edge, higher, and mostly asymmetric modes are formed, causing the sound to become harsher. The dynamic range encompasses a sound power level from 79 dB at a *pp* to 108 dB for *ff* (calculated after Burghauser and Spelda, 1971). In relation to the loudness impression for *pp*, one needs to take into consideration that at the low frequencies, relevant for the bass drum, the ear is not very sensitive at low levels, so that the bass drum can certainly reach the lower limit of an audible *pp*.

The time structure of the tone is characterized, on the one hand, by the drop in frequency of up to 140 cents, i.e., more than one half step, during the first second after the impact. On the other hand, beats of the mode pairs, mentioned earlier, in the frequency range above the fundamental, create amplitude variations, which result in the breathing character of the bass drum tone (Fletcher and Bassett, 1978). The median decay time lies around 8 s for the strongest tone contributions, in the region of 200–400 Hz it is around 4 s, at higher frequencies it drops by a factor of 2 per octave. Components below 50 Hz can ring for 15 s or longer (Plenge and Schwarz, 1967).

In addition to the felt mallet there is a second mallet, the so called "brush" used in Turkish music, it is formed from a split reed stick. It produces shorter, harder impacts of higher precision. It serves to subdivide the measures struck by the felt mallet. The symphony 100, the "Military Symphony," of Jos. Haydn is among the best known examples for this. In this symphony, for the most part, the dull felt impact marks the first beat of the measure, and the brush impacts are used for the

Score example 10 *Top:* J. Haydn, Symphony Nr. 100, 2nd movement, measure 174 *ff.* (without winds). *Bottom:* W. A. Mozart, The Abduction from the Seraglio, 3rd Act, Chorus of the Janissars (without winds and low strings)

additional beats. There are, however, places where *sforzando* impacts of the brush are found on the first beat of the measure. Mozart, in contrast, in his "Abduction from the Seraglio" lets the hard brush stroke run through as a uniform rhythm, and combines it on the first and occasionally on the third beat of the measure with the impact of the felt mallet (see score example 10).

3.7.3 Snare Drum

Since Rossini made the snare drum acceptable in his opera "La gazza ladra" in 1817, it has also found entrance into the symphony as a rhythm instrument. The rhythmic precision is accomplished by the short transient of about 7 ms (Melka, 1970), as well as by the lack of very low tonal contributions. The maximum of the radiated sound lies between 300 and 1,000 Hz, depending on the nature of the impact. For a *forte*, an impact near the middle of the membrane favors the low components, and for a *piano* an impact near the edge favors the higher partials. Both, the inharmonic location and the frequency width of the partials, prevent a unique pitch impression.

The high frequency components of the spectrum are further strengthened by the so-called snare strings, which are stretched below the low membrane. By tuning these strings below the membrane, they collide in a pulsating manner with the

membrane, and thereby excite additional high frequency vibrational modes. This increases the noise impression of the snare drum significantly (Rossing et al., 1992). The decay time of the strongest partials is of the order of magnitude of 1 s (Plenge and Schwarz, 1967), so that even a very rapid impact sequence (drum roll) is recognized as such without going over into a uniform noise. For a *pp* the sound power level lies at around 74 dB, which, because of the short duration of the individual impacts, can be perceived as very soft. For *ff* the sound power level reaches about 100 dB, which suggests a rather wide dynamic range (as calculated after Burghauser and Spelda, 1971).

3.7.4   Gong

In a number of operas with large orchestras, tuned gongs are used to create exotic tone colors. The best known example for this, – next to Saint-Saens' "La Princesse jaune" and Strauss' "Frau ohne Schatten" – could be Puccini's "Turandot," where nine gongs are required with pitches between A_2 and A_3, as well as a "Gong Grave" in A_2. Their pitch is determined by a precise fundamental and an octave partial, and occasionally supported by a double octave (4th partial). Often the pitch of this 4th partial corresponds to a seventh. The third partial is always inharmonic. It lies by a whole step to a fourth above the octave, a major third is perceived as tonally optimal. Additional higher inharmonic partials complete the spectrum, and with greater loudness a strong rushing noise is an essential part of the tone color.

While for higher pitched gongs the fundamental always dominates, in the low gongs, the 3rd and 4th partials can become by up to 8 dB stronger than the fundamental; the octave partial always lies by about 10 dB below the fundamental. The lower the gong is pitched, the louder the rush noise becomes, with an intensity maximum around 1,000 Hz, which is in the region of the vowel color of "a(ah)." Around 3,000 Hz the rush noise level for a low gong at ff is about 10 dB below the fundamental and above 3,000 Hz it drops off with 10 dB/octave, while for the higher gongs at 3,000 Hz it already lies by 20 dB below the fundamental.

The strength of the impact exerts a strong influence on the tonal picture of gongs. While the intensity of individual partials rises steadily from *pp* to *mf*, non-linear effects appear for yet stronger excitation. These lead to energy transfer between modes, whereby especially rush noise contributions become more pronounced. Thus the accessible dynamic range for a clear tone without rush noise from *pp* to *mf* or *f* only covers 17 dB with a sound power level of 91 dB at *pp* and 108 dB at *f*. A further increase by 8 dB is possible, however, that will not raise the strength of the fundamental, the additional energy appears in the higher frequencies, particularly in the 3rd and 4th partials, as well as in the rush noise. Furthermore, a consequence of these non-linear effects is that the fundamental and partial frequencies, and thus their pitch, start at up to 80 cents higher for very strong impact, and only in the course of about 4 s reach their final value. For *mf* the tone starts about 20 cents high and needs approximately 2 s to reach its final value.
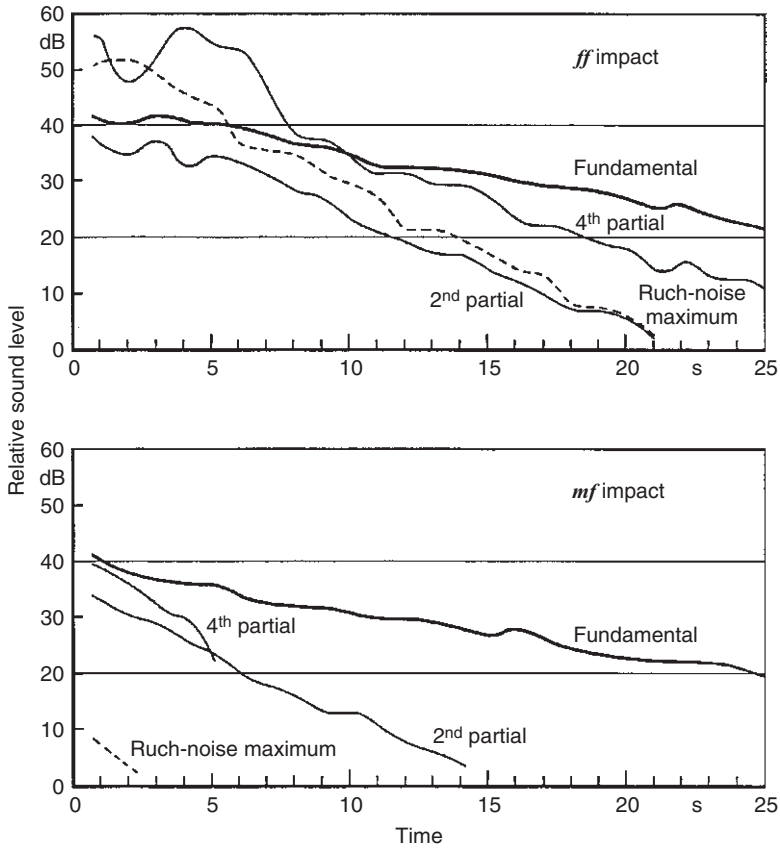
Fig. 3.29 Time dependence of the level of the most important tone contributions of a Gong Grave in A_2 for two different dynamic levels

The time dependence of the level structure for individual tonal components during the first seconds after an ff impact is particularly interesting. As Fig. 3.29 shows, the discrete partials initially become lower, while the rush noise reaches its maximal strength after 2 s, and thus exceeds all other tone contributions. After that, the rush noise decays steadily, and after 6 s becomes weaker than the fundamental, which only then becomes the strongest component. This slow spectral development endows the gong with its magnificent sound. At mf and especially at pp the individual partials exhibit a smooth level drop from the beginning. The same occurs for *ff* after a later point in time. The decay time of the fundamental is shortened from about 75 s for a Gong in A_2 to 30 s for a Gong in A_3. The decay time of the following partials, up to about five times the fundamental frequency, is also about half as long. Above that, the decay time drops uniformly down to the value of 4 s at 8,000 Hz. For "*secco-* impacts" this extremely long decay is reduced to a few seconds by damping with the free hand.

3.7.5   Cymbals

As is the case for the gong, the tonal picture of cymbals is determined in large measure by the time development of the different tonal contributions. The large number of inharmonic partials, which to some extent are densely spaced, do not permit the emergence of a pitch. At first, during the initial transient time of 10–20 ms, strong vibrations of a few radial modes are formed at frequencies around 400 Hz and also in the region between 700 and 1,000 Hz. After 50–100 ms they pass their dominant role to high frequency rush noise contributions between 3,000 and 5,000 Hz, which at times can expand to 10,000 Hz. Again, the energy exchange between different vibrational modes, or also between longitudinal and bending waves becomes a determining factor. This preferred sound radiation of high frequencies forms the tonal impression in the time frame of about 1–4 s after the impact, and results in the bright shrill sound of the cymbal. Thereafter the maximum of the sound intensity reverts back to the frequency contributions around 400 Hz. This is primarily determined by the damping of the vibrational modes, i.e., the decay behavior (Fletcher and Rossing, 1991).

The decay time of vibrational modes around 400 Hz is about 30–40 s; for components around 3,000 Hz it amounts to approximately 10 s, and for components around 6,000 Hz it is still 5 s (Plenge and Schwarz, 1967). The lowest partials in the region of 50–100 Hz, which are relatively unimportant for sound impressions in a room, can even have decay times of the order of 100 s, however, this becomes noticeable only for close microphone positions. It should be noted, that for frequencies below about 700 Hz the decrease in level initially occurs relatively rapidly, because of energy transfer to other modes, and the decay times mentioned earlier only take effect after about 200 ms, thus a time-level plot shows a break (as is the case for the piano) (Müller, 1982; Fletcher and Rossing, 1991).

Within certain limits the dynamic range of cymbals depends on the nature of the excitation. For an impact with a felt mallet, a calculation after Burghauser and Spelda (1971), yields sound power levels between 73 dB for *pp,* and 101 dB for ff. For a wooden mallet these values rise to 82 dB for *pp* and 111 dB for *ff*. Different mallets have more influence on the strength of the high frequency contributions than differences in impact. When two cymbals are crashed against each other, sound power levels between 74 dB for *pp* and 108 dB for *ff* can be expected.

3.7.6   The Triangle

Closely spaced inharmonic partials at very high frequencies play the most important role in determining the tone of a triangle. For the normal impact direction (perpendicular to the triangle plane) the spectrum reaches to over 20,000 Hz without significant level drop. The maximum of the spectral envelope is formed around 6,000 Hz. Only one partial is found below 1,500 Hz, it is located near 400 Hz. When the triangle is hit in a direction parallel to the plane, the number of

excited vibrational modes is reduced, and the partial sequence is not so dense. When, in addition to a fundamental near 400 Hz, there are two nearly harmonic partials, for example near 1,600 and 2,000 Hz, it is entirely possible that a certain pitch is noticeable in the sound of a triangle (Rossing, 1982a).

The initial transient of about 4 ms is extremely short particularly because of predominantly very high frequency contributions (Melka, 1970). For the fundamental, the decay time is about 30 s. For higher frequencies it drops rather uniformly by a factor of 2 per octave (Plenge and Schwarz, 1967). This long decay affects that for a rapid impact sequence (usually hit as a roll between two different sides of the triangle) the sound goes over into a fluctuating steady tone and assumes the character of a silvery shimmer for the overall sound of the orchestra.

The dynamic range is characterized by a sound power level of 66 dB for *pp* and 91 dB at *ff* (as calculated after Burghauser and Spelda, 1971). Because of the predominantly very high frequency contributions, the triangle is not only heard easily above the orchestra, but it is also easily located by the listener in the hall, since these high frequencies are only reflected relatively weakly by the hall.
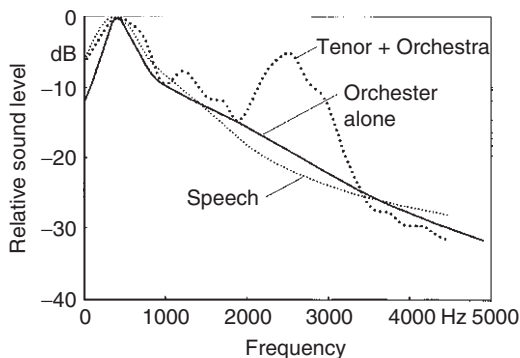
3.8 The Singing Voice

3.8.1 Sound Spectra

By nature, the spectral envelope of the singing voice is mainly characterized by formants. In each case, the sung vowel determines the location of the strongest partial. A male voice can vary the lowest formant in the region of 150–900 Hz and the second formant between 500 and 3,000 Hz. For female voices the lower limit lies higher corresponding to the tonal range. The constant change between vowels within the text consequently leads to constantly changing envelopes. Above about A_5 (880 Hz) the fundamental begins to exceed the first formant of the vowels, the tuning of the oral cavity is therefore undertaken by female singers for optimal enhancement of the fundamental and the octave (by the 2nd formant). This technique can also be used for lower voices. It serves to raise volume, however, it reduces vowel recognition, and to some extent causes tone quality to suffer. It is also used extensively with peak tones of tenor and alto voices, to lend sufficient strength (Sundberg, 1977; 1991). For tenor voices, however, generally the 2nd and 3rd partials are predominantly amplified (the fundamental less so), otherwise the voice assumes a more female character, as is the case for countertenors (Titze & Story, 1993).

The spectrum can drop strongly below the first formant. In the low register of male voices, the fundamental can lie by 15–20 dB below the strongest partials. High frequency contributions in the range between about 2,300 and slightly above 3,000 Hz have special significance. Not only is this the location of those secondary formants which make it possible to differentiate between different voices (for the same vowel), but it is also the location where for trained voices the so-called singer's formant develops. This singer's formant can provide a quality criterion

Fig. 3.30 Typical envelope for Speech, Orchestra, and singing voice with pronounced singing formant (after Sundberg, 1977)

for the singing voice (Winckel, 1971; Sundberg, 1977). It can reach a level of within 5 dB of the strongest partials. Since the orchestral instruments radiate much weaker overtones in this frequency region, the singer's formant lends to the voice an ability to carry the tone and transcend the orchestra (see Fig. 3.30). As determined by the varied lengths of vocal tracts, a typical frequency location of the singer's formant occurs around 2,300–2,500 Hz for a bass, 2,500–2,700 Hz for a baritone, and 2,700–2,900 Hz for a tenor. The typical frequency location of the singer's formant for a female voice lies somewhat higher, like 2,900 Hz for a mezzosoprano. Above 3,500 Hz the spectrum drops steeply at a rate of approximately 25 dB/octave. Exceptions lead to a timbre which is too metallic. In "belting," the vocal technique frequently used in musicals, in which, by elevating the larynx, the singer's formant is raised to above 3,500 Hz, the high frequency components are deliberately strengthened (Estill et al., 1993). It is of interest that this frequency placement finds a parallel in formants of the harpsichord (see Sect. 3.5.1). Very high frequencies are nevertheless important for recognition of consonants. Voiced sibilant sounds reach up to about 8,000 Hz, unvoiced sibilants even up to 12,000 Hz.

3.8.2   Dynamics

The dynamic range of all singing voices is characterized by a clear rise from the low register to the higher positions. As shown by the measurement results of Burghauser and Spelda (1971) – converted to sound power level – in Fig. 3.31, for solo voices the lower dynamic limit of low male voices in the lower register lies around 70 dB, for a dramatic tenor and for female voices around 60 dB, for high registers it rises to values between 85 and 110 dB. The upper dynamic limit lies between 85 and 95 dB for the low register, and it rises up to 110–125 dB in the upper region. This results in a median dynamic range of 25–30 dB, which, for favorable tone locations can be widened to more than 40 dB.

Since, once during each vibration cycle, the vocal cords close completely at higher dynamic levels, while for soft singing there is always a residual opening, the overtone content increases with increasing dynamics. This effect is very
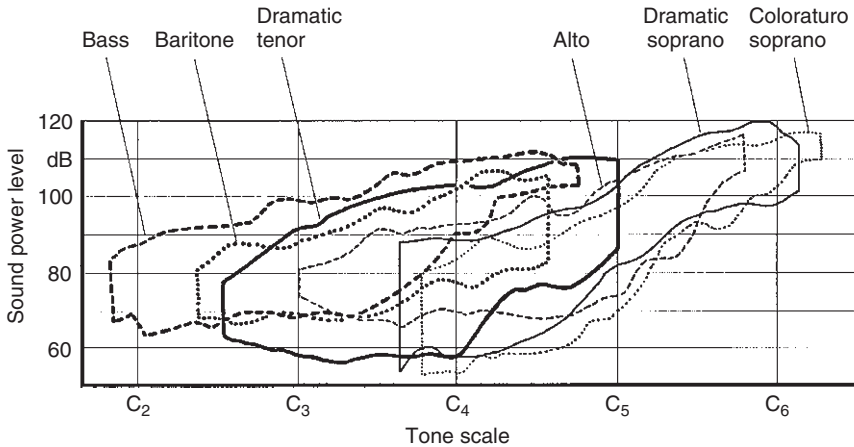
Fig. 3.31 Dynamic range of solo singing voices (after Burghauser and Spelda, 1971)

pronounced for a singer's formant which increases by about 1.5 dB for every 1 dB increase of the strongest partials. For individual voice control in a noisy environment, it is of interest to note that the sound level at the ear of the singer lies 10 dB below the radiated sound pressure level (Ternström and Sundberg, 1983).

With choral singers, the sound power level generally is not as high as for soloists, though the difference should not be that large for professional opera choruses. For certain lay choirs, Ternström (1989) found a dynamic range which led to an average sound power level for individual singers from 71 dB at *pp* to 97 dB at *ff*. For boys choirs the dynamic range of individual singers is narrower and reaches from about 80 dB at *pp* to 91 dB at *ff*. This suggests for an average *forte* an order of magnitude of 88 dB for boys and 91 dB for adult choir voices.

3.8.3 Time Structure

For a singer, the initial transient is determined by the nature of the beginning consonant. Explosive sounds lead to a very short noise impulse of from 20 to 30 ms duration; already after 40–60 ms the full harmonic sound can be developed. In contrast, sibilants are characterized by a duration of about 200 ms. For an initial "m," a 40 to 50 ms noise is immediately followed by a humming phase – for a closed mouth – lasting up to 150 ms, before the full tone is developed. It is interesting to note that the singer's formant already comes in during this humming phase. An initial "r" is characterized by a noise impulse sequence, whose individual impulses follow each other with a 35–45 ms separation.

The singers' vibrato usually moves in the frequency region of 5–7 Hz (Winckel, 1960), whereby the vibrato frequency typically increases slightly toward the end of the tone (Prame, 1993). For a *vibrato* width from about ±40 to ±80 cents, a pure frequency modulation results without change in the envelope. At most, individual
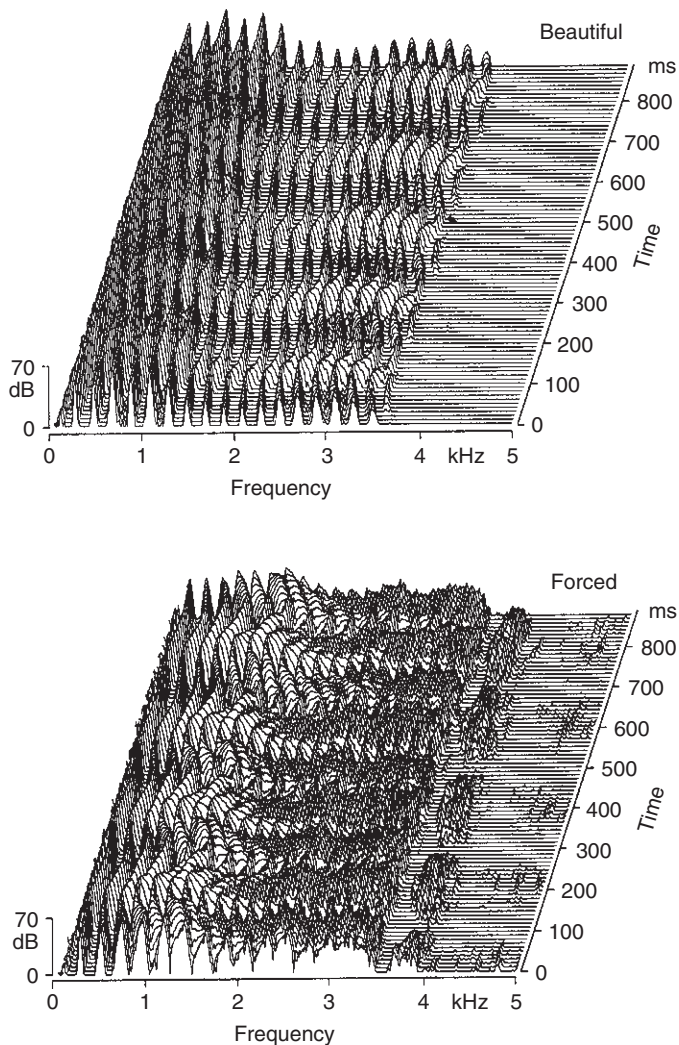
Fig. 3.32 Time variation of the spectrum for a singer's *vibrato* (Baritone, sung pitch G_2)

partials will slide up or down on the flank of the formant resonance curve, which will actually contribute to the clarification of the formant (Benade, 1976). Nevertheless, under those circumstances the tone color remains constant in time. For a very forced vibrato, as shown in Fig. 3.32, it is a different matter. This vibrato, which – not only by reason of its width of more than ±200 cents – is sensed as esthetically unsatisfying, is additionally characterized by a constant phase amplitude modulation with high tonal and noise contributions. This makes the tone especially noticeable, if not even penetrating. Depending on musical context, a

vibrato tending in that direction can occasionally be very appropriate, to increase the voices power to stand out.

3.8.4 Choral Singing

Fusing numerous voices into a homogeneous choral sound demands not only mutual adaptation on the part of the singers, but also the avoidance of all effects which make an individual voice stand out. Therefore, the singer's formant, so essential for the soloist, becomes an annoyance for choral singers, unless it is present in comparable strength in all individual voices. In professional choirs it is approximately 5–15 dB weaker than in soloists (Sundberg, 1990), while it is almost not found at all for lay choirs. As a result, individual voices in a lay choir, which emphasize the singer's formant, stand out in the overall sound, furthermore, an additional vibrato is particularly dangerous. In lay choirs, generally the 3,000 Hz (at *mf*) components are from 20 to 25 dB weaker than the strongest tonal contributions. This can reach 30 dB for boys' choirs (Ternström, 1991b).
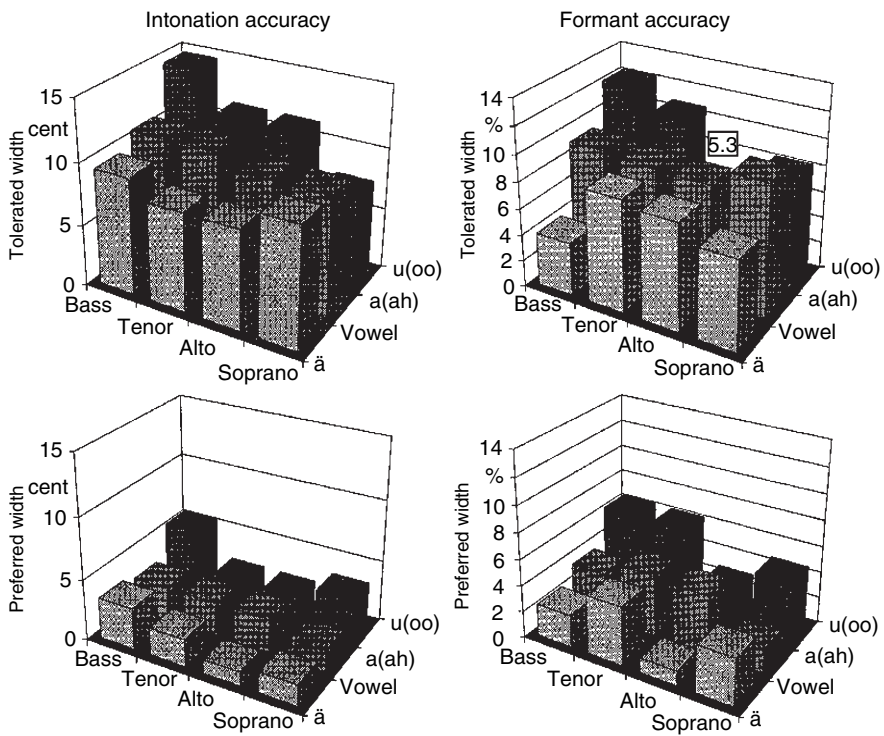

Fig. 3.33 Subjective evaluation of intonation accuracy and formant determination (after Ternström, 1991a)

Basic for an effective all-inclusive choral sound is the accurate intonation of all singers. In hearing tests, Ternström (1991a), using synthetic choral sounds, determined what accuracy is preferred, or alternately is tolerated. The results are represented in Fig. 3.33 for each of four voices and three vowel colors. Accordingly, intonation is considered good, if the standard deviation (width of intonation spread) clearly lies below ± 5 cents (i.e., two thirds of the voices should fall within theses limits). An intonation accuracy of ± 10 cents is still tolerable. Bass voices are relatively insensitive to these limits, particularly for dark vowels. In contrast Lottermoser and Meyer (1960) found significantly larger deviations. An extreme case is given by the Don Kossaks with up to ± 60 cents. Tuning the formant frequencies, i.e., the matching of vowel sounds between individual singers of a choir, is also important. As indicated in the pictures on the right of Fig. 3.33 for the lower two formants, efforts should be made to keep the standard deviations significantly below $\pm 6\%$, while $\pm 9\%$ can still be tolerated. For the higher formants $\pm 12\%$, should be maintained, a standard which can be reached only by very systematic choral training.