Dept. for Speech, Music and Hearing

Quarterly Progress and Status Report

Measuring inharmonicity through pitch extraction

Galembo, A. and Askenfelt, A.

journal: STL-QPSR
volume: 35
number: 1
year: 1994
pages: 135-144

Measuring inharmonicity through pitch extraction[1]

Alexander Galembo[2] and Anders Askenfelt

Abstract

Some experiments are described which illustrate new applications of pitch extraction processing. It is shown that cepstrum *and* harmonic product spectrum *contain distinctive information about the inharmonicity of musical sounds - in this case piano tones - and that they can be successfully used for illustration of the phenomenon. A way is proposed how to improve the methods, enabling a direct quantitative measurement of inharmonicity in tones from musical instruments.*

Introduction

Due to the stiffness of real piano strings, the partial frequencies are raised a little compared to the harmonic multiples of the fundamental frequency of the ideal string. The deviation is well described by the classical expression (Fletcher, 1964):

$$f_n = n\, f_0\ (1+n^2B)^{0.5} \qquad (1)$$

where f_n is the frequency of the n-th partial, f_0 the fundamental frequency of the ideal string, and B the inharmonicity coefficient. For typical piano strings, B ranges between 0.00005 and 0.0007 approximately, in the lower half of the instrument's compass.

This "well-behaved" inharmonicity is a hidden parameter in the tone, which is not possible to observe directly in waveforms or spectra. In fact, up to now, inharmonicity has not had any graphical illustration by direct measurement. The normal procedure of measuring inharmonicity includes spectral analysis, (manual) identification of the partials, measurement of the partial frequencies, and calculation of the inharmonicity coefficient B from Eq. (1), preferably by comparison of several pairs of partial frequencies and averaging.

Some of the pitch extraction procedures developed in the early days of digital speech processing seem to offer a way to overcome these difficulties. In particular *cepstrum* and *harmonic product spectrum* show interesting features, which enable a direct estimation of the inharmonicity coefficient. The idea of analysing inharmonicity by using pitch extraction algorithms was tried for the first time rather long ago by the first author in the acoustical laboratory of the piano factory "Red October" in Leningrad (Galembo, 1979; 1986), but could not be carried further at that time.

Perceptually, inharmonicity is related to how clearly the pitch will be perceived. A harmonic spectrum will evoke a more pronounced sensation of pitch than a strongly inharmonic spectrum. Terms like "pitch strength" or "pitchiness" have been used to characterise this quality (Rakowski, 1977; Fastl & Stoll, 1979).

---

[1] This paper was presented at the Stockholm Music Acoustics Conference (SMAC), July 28-Aug. 1, 1993

[2] 47 Furstadtskaia St., Apt. 17, 191123 S:t Petersburg, RUSSIA

Cepstrum and harmonic product spectrum

CEPSTRUM, which today is a standard processing in many speech applications, describes the periodicity of the spectrum (see e.g. Noll, 1967). The standard definition is given as "the power spectrum of the logarithm of the power spectrum":

$$C(\tau)=|\mathcal{F}(\log|F(\omega)|^{2})|^{2} \qquad (2)$$

where F (ω) is the spectrum of the measured signal.

The result of a cepstral analysis of a strictly periodic (harmonic) sound is in principle a δ-function, located on the quefrency (time) scale at the fundamental period of the signal (Fig. 1). Inharmonic sound spectra, as those described by Eq. 1 with a gradual increase of the distance between adjacent partials, will broaden the cepstrum peak, or even give a region of peaks. If *n* partials with raised frequencies according to Eq. 1 are included in the analysis, it can be shown that the cepstrum peak will occupy a region of width

$$\frac{w_{CEPS}}{\tau_o} = \frac{n(1+n^2B)^{0.5} - (n-1)[1+(n-1)^2B]^{0.5} - (1+B)^{0.5}}{(1+B)^{0.5}\,\{n(1+n^2B)^{0.5} - (n-1)[1+(n-1)^2B]^{0.5}\}}$$

normalised by the fundamental period of the ideal string $\tau_o = 1/f_0$. The width increases smoothly with increasing *B*.

This region, the width of which depends only on *n* and *B*, defines the pitch character of the inharmonic sound. For this reason it was coined as the *pitch peak region*. In principle, *B* can be determined directly from a measurement of the peak width, but unfortunately the exact number of partials included in the cepstral analysis is seldom known in practice. Even an elimination of the partials above a certain frequency in F(ω) will still leave an uncertainty, as the actual *n* covered by a given frequency range is dependent on the (unknown) value of *B*. In practise, however, the uncertainty is usually not more than ±1 partial for typical piano tones.

The HARMONIC PRODUCT SPECTRUM, which is a less well known procedure, is based on the fact that the partial peaks in the spectrum of a periodic signal are located at harmonic multiplies of the fundamental frequency (Noll, 1969). If a set of new spectra is generated by compressing the original spectrum along the frequency axis by successively dividing by integers 1, 2, 3, 4, . . . *m*, and then adding the compressed versions (in logarithmic amplitude scale), the result is an enhanced peak at the frequency of the fundamental, which coincides for all compressed versions (Fig. 2):

$$\log \pi(\omega) = \sum_{k=1}^{m} \log |F(k\omega)|^{2} = \log \prod_{k=1}^{m} |F(k\omega)|^{2} \qquad (4)$$

which after antilogarithming gives the harmonic product spectrum (HPS):

$$\pi_{HPS}(\omega) = \prod_{k=1}^{m} |F(k\omega)|^2 \qquad (5)$$

where m is referred to as the order of the harmonic product spectrum, indicating the number of compressed spectra (and highest divisor) included.

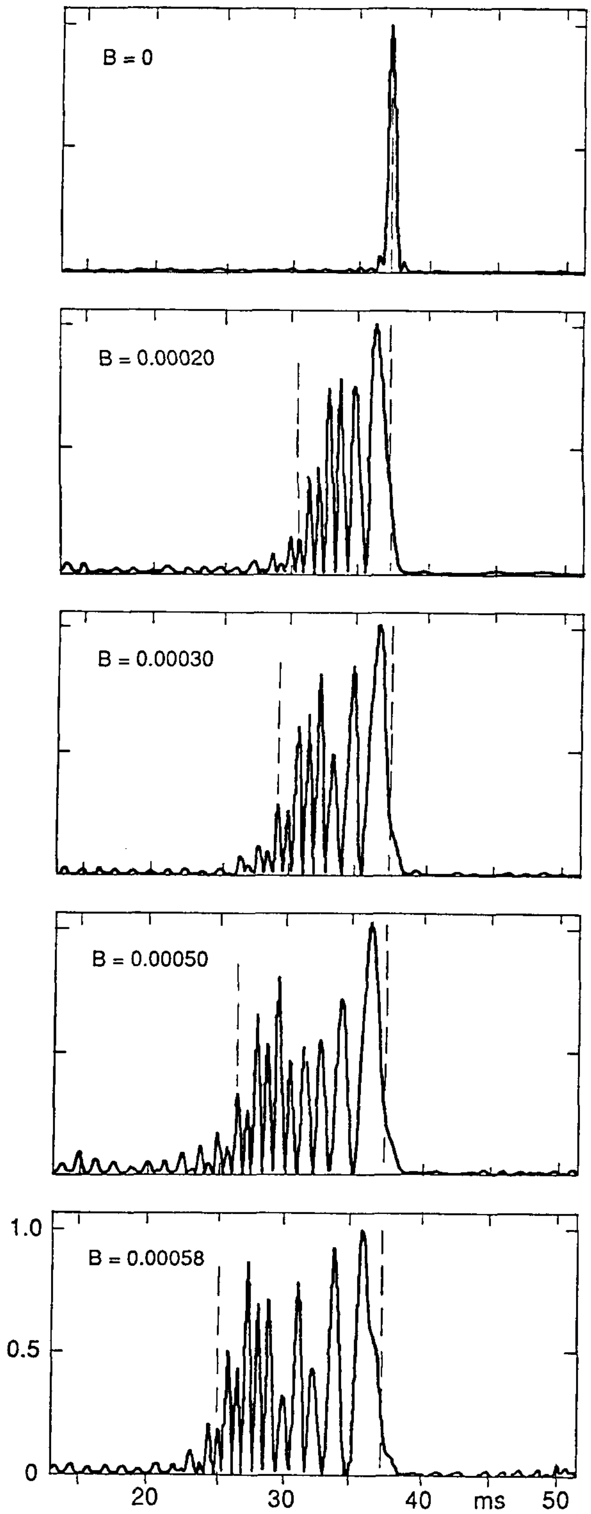

Fig. 1. Cepstrum analysis of synthesised inharmonic tones including 30 partials of equal amplitude (A_0 = 27.16 Hz) with inharmonicity coefficient B = *0, 0.00020, 0.00030, 0.00050, and 0.00058. The dashed vertical lines define the predicted pitch peak region according to Eq. 3.*

As for the cepstrum, a slightly inharmonic input spectrum F(ω) will introduce a broadening of the peaks in the harmonic product spectrum, including the dominating peak at the fundamental frequency, also in this case referred to as *pitch peak*. The relative width of the pitch peak region (normalised to f_0) will depend on B and m as

$$\frac{w_{HPS}}{f_0} = (1+m^2 B)^{0.5} - (1+B)^{0.5} \qquad (6)$$

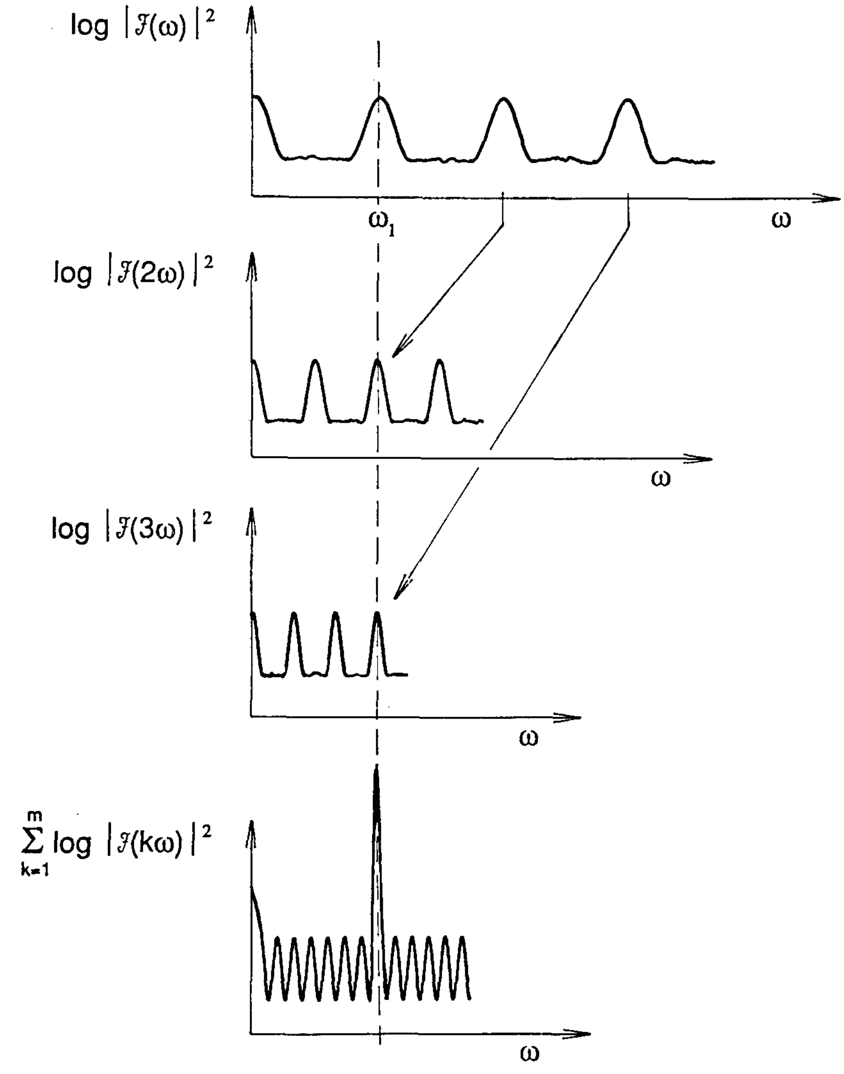

Fig. 2. Explanation of the principle of harmonic product spectrum (Adapted from Noll, 1969).

Measurements and results

The cepstrum and harmonic product spectrum methods were tested using tones from five pianos with different amount of inharmonicity. The inharmonicity coefficients of the recorded notes were first measured in the normal manner by spectral analysis and measurement of each partial frequency, followed by a matching of f_0 and B (Eq. 1) to the partial series (Table 1).

	Steinway D	*Steinway C*	*Nordiska 1*	*Nordiska 2*	*Straud*
model	concert grand	salon grand	upright	upright	upright
size	274 cm	227 cm	114 cm	114 cm	139 cm
A_0	0.000160	0.000190	0.000340	0.000350	0.000580
E_1	0.000067	0.000095	0.000200	0.000200	0.000260
A_1	0.000059	0.000093	0.000145	0.000150	0.000220
E_2	0.000070	0.000110	0.000140	0.000140	0.000220
A_2	0.000095	0.000130	0.000160	0.000150	0.000220

Table 1. Inharmonicity coefficient B *of piano five notes from five pianos with different amountof inharmonicity (concert grand - small upright).*

In addition, a series of synthetic notes with 30 spectral components of equal amplitudes and inharmonicity coefficients between 0 and 0.0005 was used to evaluate the methods. The reason was that spectra of real piano tones contain components which not at all are related to the fundamental. In particular the longitudinal string modes can be rather prominent. Such interleaved spectral components confuses the picture before sufficient experience of the analysis methods has been gained. Further, the spectral envelope, which was found to have a rather significant influence on the result of the cepstrum analysis, could be controlled in these examples.

Cepstrum

A cepstrum analysis of five synthetic 30-component inharmonic tones (A_0 = 27 Hz) with B ranging from 0 to 0.00058 is shown in Fig. 1, clearly illustrating the broadening of the pitch peak region with increasing inharmonicity. Also the fundamental peak is seen to raise in frequency (shifting to the left on the quefrency scale), with increasing inharmonicity. The results agree well with predictions in the sense that the peak region is closely defined by the limits calculated according to Eq. 3.

The corresponding cepstral analysis of the lowest note ($A_0 \approx$ 27 Hz) for the five pianos are shown in Fig. 3. Again, the pitch peak region is seen to be well described by the calculated limits. It is obvious from Figs. 1 and 3, that the cepstral analysis gives a good illustration of the degree of inharmonicity, in satisfactory agreement with the theoretical predictions.

An interesting perceptual comparison is included in the upper part of the cepstrums in Fig. 3, showing the upper and lower (quefrency) limits for the fundamental of a harmonic spectrum (from a generator), matched to the pitch of the real piano tone by two experienced piano tuners. As seen, the tuners consistently perceived a pitch in the lower (right) end of the pitch peak region. The subjects considered the task as difficult, and the width of the matched interval was also relatively large, approximately 50 ¢ for this low note. A little unexpected, a clear influence of the inharmonicity on the width of the matched interval was not observed, in spite of the fact that B differed roughly a factor four between the extremes.

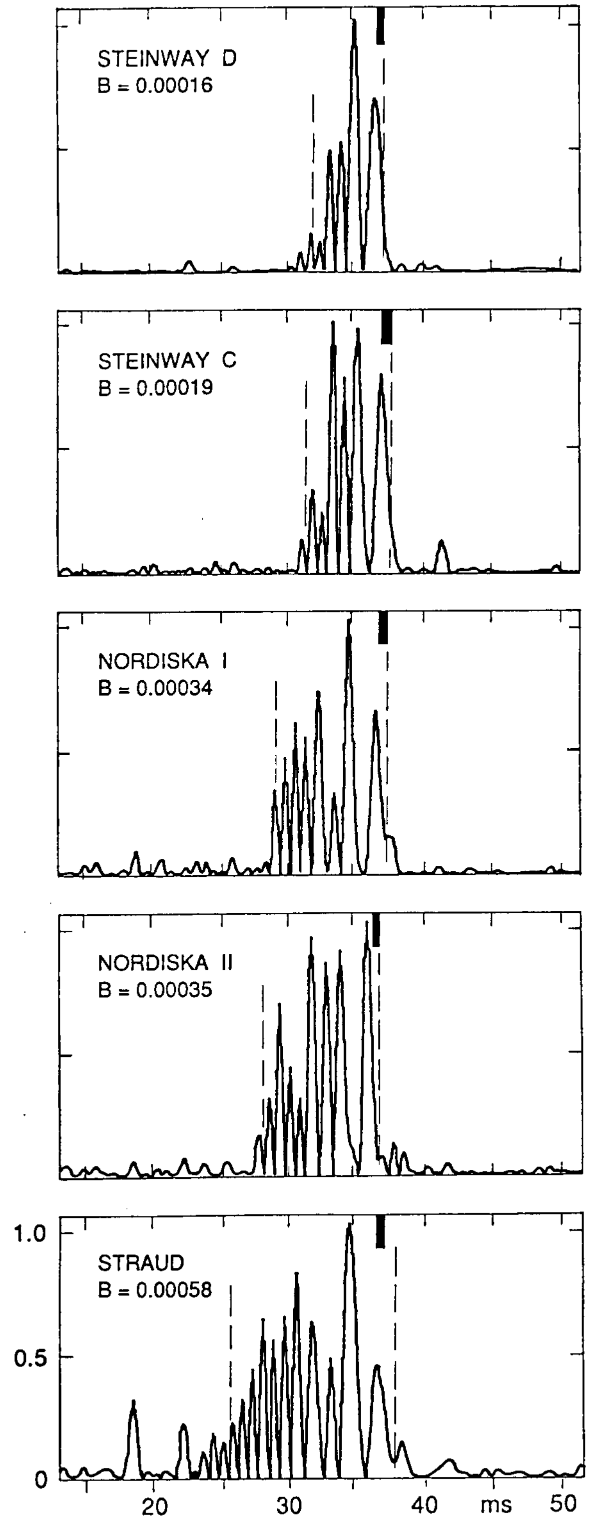

Fig. 3. Cepstrum analysis of real piano tones from five different pianos ($A_0 \approx$ 27 Hz) with inharmonicity coefficient B = *0.00016, 0.00019, 0.00034, 0.00035, and 0.00058. The dashed vertical lines define the predicted pitch peak region according to Eq. 3. The filled bar at the top of each plot indicate the perceived pitch in a matching experiment with harmonic spectra (see text).*

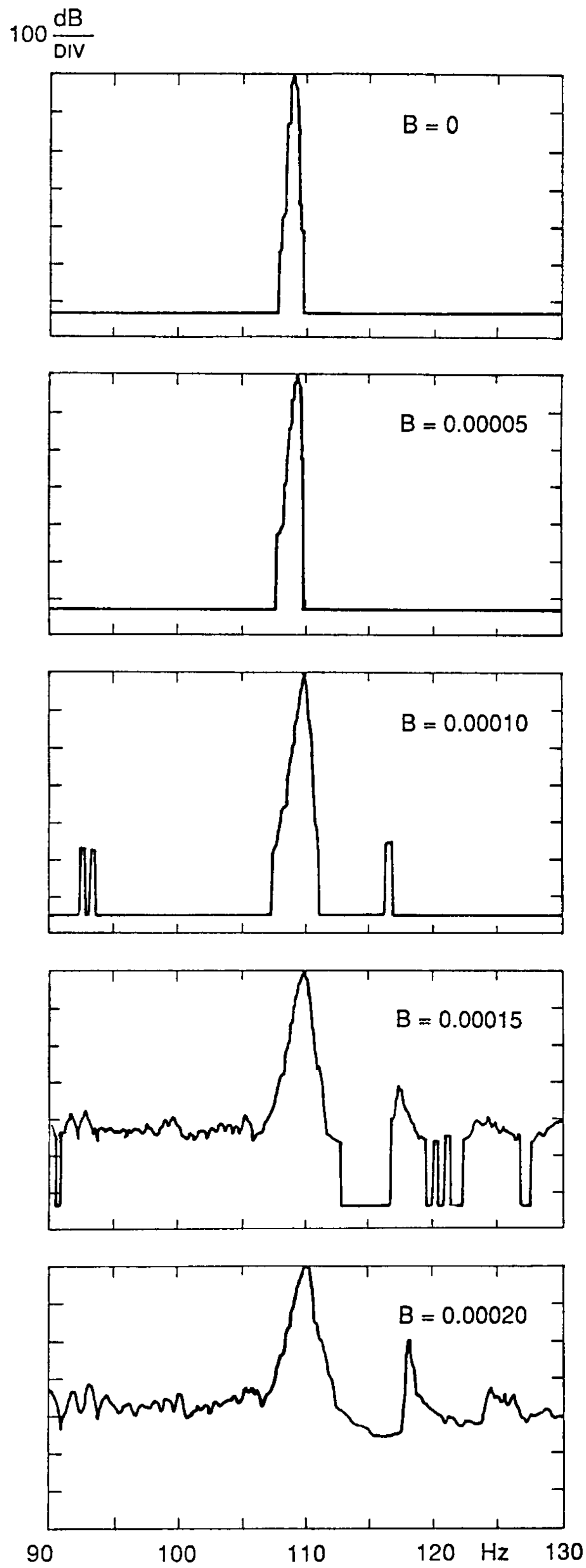

Fig. 4. Harmonic product spectrum of synthesized inharmonic tones ($A_2 \approx$ 110 Hz) including 30 partials of equal amplitude with B = *0, 0.00005, 0.00010, 0.00015, and 0.00020. The number of compressed spectra included in the harmonic product spectrum is* m = *18. Note the amplitude scale which covers 700 dB!*

Harmonic product spectrum

A harmonic product spectrum analysis of five synthesised tones ($A_2 = 110$ Hz) with B ranging from 0 to 0.00020 is shown in Fig. 4. The pitch peak is seen to rise prominently over the backgroundlevel for low inharmonicity, and decrease and broaden with increasing B. The corresponding analysis for the same note (A_2) from two different pianos is shown in Fig. 5. Both figures show the expected dependence between inharmonicity coefficient B and width of the pitch peak. As can be seen in Fig. 5, the pitch peak region is reasonably well defined by the predicted limits according to Eq. 6. It could also be observed that the pitch peak is not so distinct for high B values, rising less prominently above the background level. This behaviour is to be expected due to the following reason.

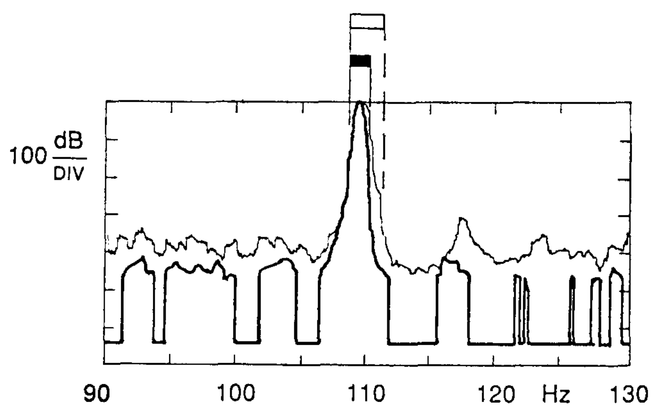

Fig. 5. Harmonic product spectrum ($m = 18$) of real piano tones ($A_2 \approx 110$ Hz) from two instruments; concert grand, Steinway D, $B = 0.00010$ (broad line), and a small upright, Straud, $B = 0.00016$ (thin line). The horizontal bars at the top indicate the predicted width of the pitch peak regions according to Eq. 6.

As mentioned, the harmonic product spectrum of a strictly harmonic sound is characterised by an exact coincidence of the frequencies of all the compressed partials of f_k/k-type, and the result is high peak at f_0. If the sound is slightly inharmonic, these frequencies are not exactly matching, though still close. This means that the pitch peak broadens with increasing B, which is the feature the method is based on.

However, the closest adjoining components in the compressed spectra, (f_{k+1}/k) and (f_{k-1}/k), move successively closer to f_0 (f_k/k) as k is increased, and depending on the analysis bandwidth they may very well merge with the pitch peak, when adding up the harmonic product spectrum. This gives an undesirable "distortion" of the pitch peak, in the sense that its true width can not be determined in a reliable way.

The situation is illustrated in Fig. 6, which shows the calculated frequencies of the compressed partials for three cases with increasing degree of inharmonicity. As seen in the figure, a separation of the pitch peak is always possible for a harmonic spectrum for any realistic m, while an inharmonic input spectrum requires a limitation of m in order to give a meaningful result. A closer inspection suggests that for low inharmonicity a reasonable value of m would be between 15 and 20. For high values of B several of the compressed spectra will inevitabely merge as soon as m exceeds 10 - 12, making a precise determination of the width of the pitch peak hazardous (Fig. 6c).

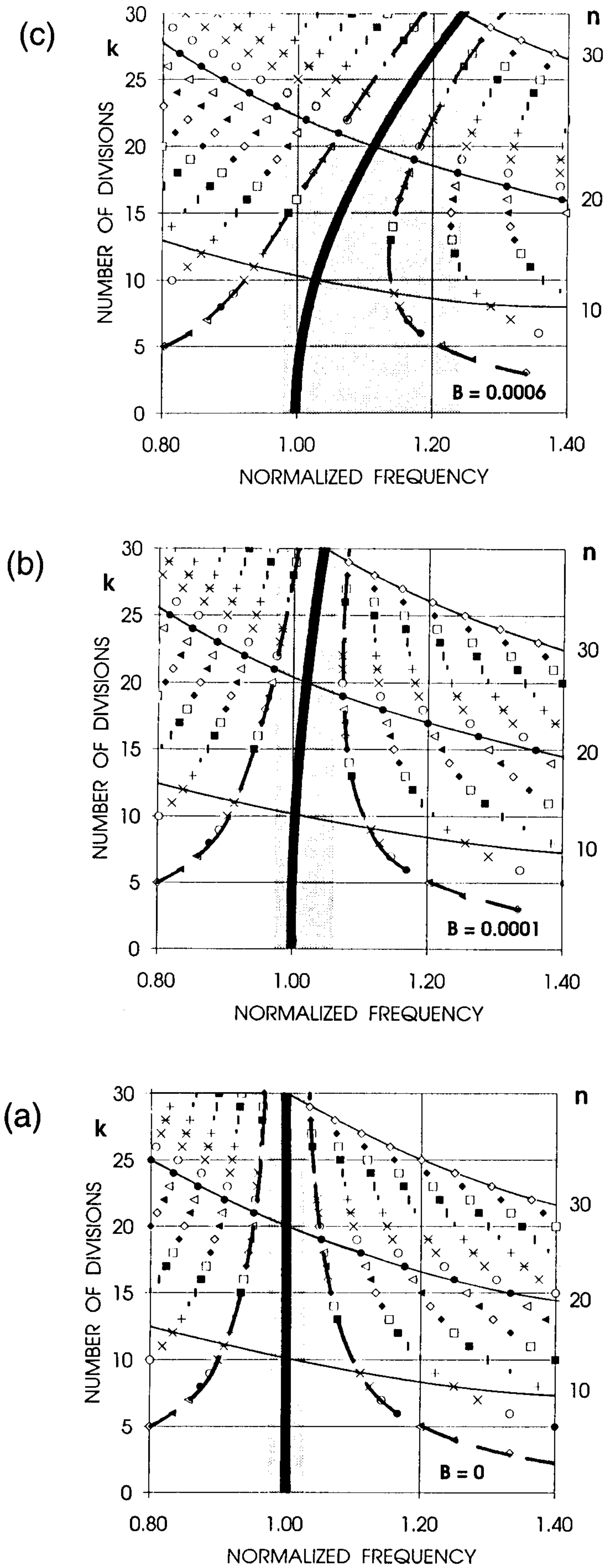

Fig. 6. Illustration of the limitations in separating the pitch peak in a harmonic product spectrum. The figures show calculated center frequencies of the partials in compressed spectra (normalized to f_o) for different values of the divisor k. For a harmonic product spectrum of order m all partial frequencies below a horizontal line m = k will be added (making the number of partials summarized in the pitch peak equal to m).
(a) Harmonic input spectrum (B = 0),
(b) slightly inharmonic spectrum (B = 0.00010),
(c) strongly inharmonic spectrum (B = 0.00060).
The curves show: f_k/kf_o (full line), f_{k-1}/kf_o and f_{k+1}/kf_o (dashed). The different markers represent partials n = 1 through 30. A separation of the pitch peak region (shaded area) from adjacent compressed partials is easy for harmonic sounds but hazardeous for high values of B and m.

Discussion

The examples presented in Figs. 3 and 5 show that either cepstrum analysis or harmonic product spectrum can be conveniently used for illustrating inharmonicity in musical sounds by a direct measurement. The analysis of the synthesised tones with known inharmonicity shows also that the measured width of the pitch peaks are in satisfactory agreement with predictions (Eqs. 3 and 6), which leads to a primitive way of determining *B*. As mentioned, however, several factors in the processing will decrease the exactness in the determination of *B* by this method.

An appealing alternative would be to include an optimising method for the inharmonicity determination. The essence of this strategy is to find the maximum amplitude of the pitch peak by variation in a parameter (inspired by the *maximum likelihood estimation* procedure for pitch extraction, Noll, 1969). This is done by introducing an inharmonicity scaling factor $k(1+k^2b)^{0.5}$ in the calculation of the compressed spectra which form the harmonic product spectrum (which now would be more appropriate to rename as "inharmonic product spectrum"). The partial frequencies are then compressed slightly more than the nominal ratios 1, 2, 3, . . . The idea of adapting the harmonic product spectrum to inharmonic sounds by using a non-integer divisor was proposed already by Noll (1969). Following Eqs. (1) and (5), the inharmonic product spectrum is now given by

$$\pi_{IHPS}(\omega) = \prod_{k=1}^{m} \left| F[k(1+k^2b)^{0.5}\omega] \right|^2 \qquad (7)$$

By variation of *b* in the expected region for the inharmonicity coefficient *B*, a maximal amplitude of the pitch peak will be reached when *b* equals *B* of the measured sound. In the lower half of the piano compass, a variation range of $0 < b < 0.0007$ would be reasonable. The frequency coordinate of the pitch peak for the optimal value of *b* is equal to the actual fundamental frequency.

By this method, the need for separating the pitch peak from from neighbouring components is circumvented, which would increase the usefulness of the harmonic spectrum technique for a precise determination of the inharmonicity. An equally appealing feature is that this method offers the possibility of an automatic measurement of the inharmonicity coefficient, in principle even in real time.

Conclusions

The experiments have shown that cepstrum and harmonic product spectrum processings give an informative graphical representation of the inharmonicity of musical sounds, which in turn enables a calculation of the inharmonicity coefficient. As such, they may be used as new, effective methods for illustration, comparison, and evaluation of inharmonicity. The harmonic product spectrum could be improved by implementing an algorithm which uses a parameter scaling algorithm, allowing for an automatic calculation of the inharmonicity coefficient of sounds from musical instruments.

Acknowledgements

The authors are indebted to Johan Liljencrants for valuable discussions on the pitch processing procedures and indispensable help in programming, to Prof. Andrzej Rakowski, Chopin Academy of Music, Warsaw for discussions on the perceptual aspects on inharmonicity, and to piano technicians Hans Norén, Swedish Radio Company, and Conny Carlsson, Royal Swedish Academy of Music, for participating in the experiments. This study was made possible by a guest research grant from the Swedish Institute to the first author.

References

Fastl, H. & Stoll, G. (1979). "Scaling of pitch strength," *Hearing Research* 1, pp. 293-301.

Fletcher, H. (1964). "Normal vibration frequencies of a stiff piano string," *J. Acoust. Soc. Am.* 36, pp. 203-209.

Galembo, A. (1979). "Method for objective evaluation of the sound quality in piano trebles," (in Russian). *Proc. of 18th Acoust. Conf. of Czechoslovakia.* Also: Galembo, A. (1981). "Electronic device for the evaluation of sounds of musical instruments", (in Russian). *USSR Invention Certificate* 828215, Int.Cl. G10 1/02.

Galembo, A. (1986). *The Piano - quality of sounds.* Moscow, 1986 (in Russian), p.16.

Noll, A.M. (1967). "Cepstrum pitch determination," *J. Acoust. Soc. Am.* 41, pp. 296-309.

Noll, A.M. (1969). "Pitch determination of human speech by the harmonic product spectrum; the harmonic sum spectrum, and a maximum likelihood estimate," *Proc. of the Symposium on Computer Processing in Communications,* Polytechn. Inst. of Brooklyn, pp. 779-797.

Rakowski, A. (1977). "Measurements of pitch," *Catgut Society Newsletter* No: 27, pp. 9-11.