ISO 532 – Living and working with alternative loudness standards

Joachim SCHEUREN

Müller-BBM GmbH, Germany

ABSTRACT

By its very nature, standardization claims to set unique rules and procedures, thus aiming unique, i.e. comparable and reproducible results. Also, to be effective, the standard and the underlying approach need to be broadly accepted by the addressed user community. To come up with such general agreement it is typically required that a “best” approach, covering all relevant applications of interest, has been established. If this does not apply, the existence of alternative standards in parallel may offer a way out.

The revision process of ISO 532 “Acoustics – Methods for calculating loudness” showed that it was not possible to prove and agree on a single “best” approach. Acknowledging the legitimate demand of applicative continuity, the member bodies of ISO TC 43 “Acoustics” therefore decided to maintain the given situation by specifying two alternative approaches again. With reference to existing national standards (DIN 45631/A1 and ANSI S3.4), these are the Zwicker approach in an extended version for time-varying sounds and the Moore/Glasberg approach for stationary sounds at first.

Based on a review of the revision process and a summarizing justification of the resulting decision, the discussion of practical and theoretical consequences of the renewed dualism will outline the chances for broader applications of loudness calculations. However, this dualism also establishes the commitment to thoroughly identify and seize the chance for final unification of the standard.

Keywords: Loudness, ISO-Standard I-INCE Classification of Subjects Numbers: 63.1; 81.2.1

1. INTRODUCTION

Since almost 40 years psychoacoustic loudness evaluations for stationary sounds can be based on an international standard, ISO 532, which, since 1975, offers two alternative methods to do so: a method A based on an approach of S.S. Stevens as defined in ANSI S3.4 and a method B based on an approach of E. Zwicker as defined in DIN 45631:1967. It took many years until this basic parameter of subjective hearing impression gained some applicative acceptance and thus started to pioneer the concept of psychoacoustic metrics at all. Being mainly due to applicative restraints together with effort related concerns, this reserve did not seem to be caused by the availability of two standardized methods in parallel. On the contrary, instead of being a handicap, these alternate approaches possibly encouraged tentative, experimental use of the loudness concept in practice even.

Nevertheless, with the decades running off, it appeared that method A continuously lost practical attention while method B - at the same time – was able to get more and more established as a proven standard for loudness evaluations in practice. Thus, initial scientific uncertainty could be overcome by verification of practical use. However, although proven, this standard together with the application of the loudness concept at all remained sensitive in the sense of still lacking wide and undisputed acknowledgement in practice.

In the end, it can be stated that - for stationary sounds - the existence of standardized approaches to estimate the subjective hearing impression “loudness” by calculation from measured data strongly encouraged the use of this concept in practice. This benefit of an existing standard clearly exceeded the handicap of methodological (A and B) alternatives. Of course, providing two standards with differing results concedes some scientific uncertainty. But any confusion caused by this ambiguity was overcome by the practical value of having available and applying a standard at all.

It thus seems to be evident that usefulness of a standard may overcome its ambiguity in a self-healing process and thus may have higher priority.

2. RETROSPECTIVE SURVEY

Although ISO 532 was kept unchanged since its first introduction in 1975, the corresponding normative scenario has been changed three times in the decades since then.

2.1 Updates of Zwicker Approach

A first updating of the Zwicker method was implemented in Germany by revision of previous DIN 45631 (from 1967) in 1991 ([2]). This new version improved the one before by specifying corrections to match the ISO equal loudness contours of that time (ISO 226:1987) for low frequencies. Being necessary to enable a software implementation of the standard, this specification was in the frame of user margins for low frequencies as previously given within ISO 532 B, however. Thus being a particular implementation of ISO 532 B only, the DIN version could not cause any critical discontinuity in data obtained so far. Also, if such discontinuity within the margin of ISO 532 B would have played a roll at all, it was overcome in the years thereafter.

A second and last updating of the Zwicker method was introduced as an appendix DIN 45631/A1 to DIN 45631 in 2010 ([3]). This appendix extends the method to arbitrary non-stationary sounds without any changes of the stationary approach itself, thus preserving and guaranteeing full compatibility with the previous version for stationary sounds (stationary sounds included as special cases).

Further adaptations have been retained, thus giving preference to the immense value of normative continuity against following some improved procedural insight or new regulations like the equal loudness contours of ISO 226:2003. Although being subject of temporary doubts, this strategy of continuity has been widely acknowledged by the user community so far.

It can be concluded that the DIN45631/A1:2010 version of the Zwicker method has not modified but precisely specified and extended the ISO 532 B standard only to enable unique algorithmic formulation in a computer program and application to arbitrary, stationary and non-stationary sounds.

2.2 Introduction of Moore/Glasberg Approach

The ongoing loss of practical relevance of the Stevens approach was given normative attention when considering revisions of ANSI S3.4. In 2005, this led to the complete replacement of the Stevens method by introduction of the new Moore/Glasberg approach in ANSI S3.4 (revised in 2007, [4]). It was most natural, of course, that this new method was matched to the new equal loudness contours as specified in the revised ISO 226:2003 standard. Also, by applying differing concepts, the method was able to reduce procedural compromises like slight unsteadiness risks between neighbouring frequency bands ([5]).

2.3 Consequences for ISO 532

With ISO 532 B having lost its applicative basis and thus turned to practical insignificance over the years, the relevance of ISO 532 was completely restricted to the relevance of its method B, thus turning it to a proven and frequently used reference. However, although serving as a reference, for the most part it was not ISO 532B:1975 which defined the underlying calculation procedure. This was and is due to the lack of numerically complete specifications of the calculative procedure within ISO 532B which, in 1975, had given preference to graphical data.

Therefore, in practice, using the Zwicker method with reference to ISO 532B typically meant applying the Zwicker method as defined in DIN 45631. Since the 1980s all leading software suppliers had implemented the 1991 version of DIN 45631 which later was replaced by the version for non-stationary sounds - long before their implementations were used to define the extending standard DIN 45631/A1 ([3]. Thus, ISO 532 or it's method B, respectively, were fully identified with extended DIN 45631/A1:2010 in the end.

Of course, beyond ISO 532 and DIN 45631, the new ANSI S3.4 approach was able to gain ground in science and in some application areas like audiology, for instance. It was therefore obvious to start revision of ISO 532 to adapt it to methodological and practical reality in science and industry.

3. REVISION OF ISO 532

Based on a decision of ISO TC 43 "Acoustics", its WG 9 "Method for calculating loudness level" started to work on the revision of ISO 532 in 2007. First milestones of this work have been

- questionnaire based survey to all ISO TC 43 member bodies to explore some key positions,
- subsequent decision to maintain a separate standard for stationary sounds,

- subsequent decision to include two methods, the Zwicker (according DIN 45631) and the Moore/Glasberg approach (according ANSI S 3.4) into new ISO 532, thus maintaining the twin status of the standard,
- difficulties in quantitatively correlating the results of different approaches and their respective uncertainties for a sufficient set of practical sounds.

Unfortunately, discussions around the latter difficulties were increasingly characterized by procedural egoism. While trying to use these difficulties as justification for standardizing one approach only, argumentative positions were taken which strongly contradicted any consensus requirement.

The related discussions mainly were about two fields of interest frequently conflicting with each other: a rather dynamic request to follow any new procedural insight and a more static demand for continuity to allow for long term comparisons of sound characteristics. To give some insight into the pros and cons, the following section roughly compares the two approaches with respect to the basic arguments of the standard discussions. More detailed definitions and explanations can be found in the respective literature ([6]).

4. COMPARISON OF METHODS

The approaches being considered here differ in concept. While the Method of Zwicker can be seen to be phenomenological in the sense of describing and putting together relevant hearing phenomena like spectral and temporal masking in an algorithm, the method of Moore/Glasberg models the hearing sensation in the sense of consecutive transformations corresponding to physiological and psycho-physical characteristics of hearing.

It is not surprising that the difference in concept is linked to differences in results when applying the methods to sounds. Obviously, any attempt to systematically evaluate or even correlate such results is strongly complicated by the unlimited variety of differing sounds and sound characteristics. To illustrate this, two comparisons shall be reviewed here,

- a comparison for sinusoidal sounds as used for the evaluation of equal loudness contours, and
- a comparison for pink noise.

As can be seen from figures 1 and 2, ANSI S3.4:2007 better matches the new ISO 226:2003 contours defined in 2003 than DIN 45631:1991 ([7]). This is evident because the respective DIN approach had been introduced 12 years before the respective ISO 226 contours. Consequently, they better match the previous ISO 226 contours defined in 1987.

As purely sinusoidal sounds represent a particular class of sounds only, their use in being representative for a broader class of sounds including real life sounds is very limited. Also, it should be kept in mind that serious doubts had been brought forward from the very beginning against poorer low frequency weightings by the new ISO 226 curves when revising their predecessors from ISO 226:1987 ([8].

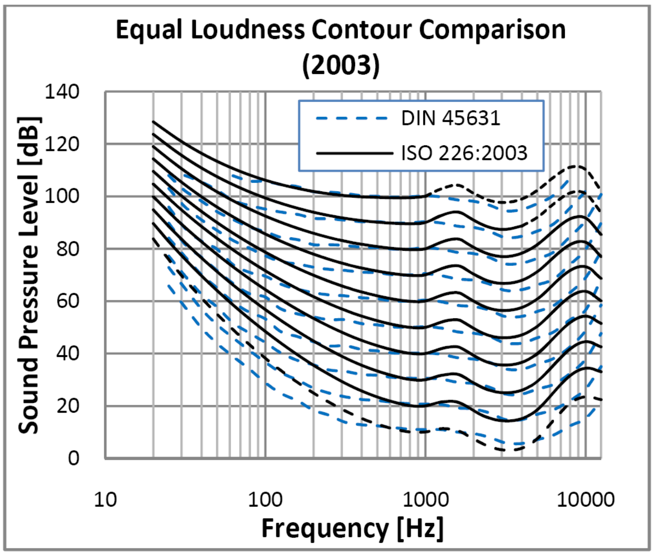

Figure 1 – Comparison of equal loudness contours as obtained from DIN 45631:1991 with ISO 226:2003 curves ([7]).

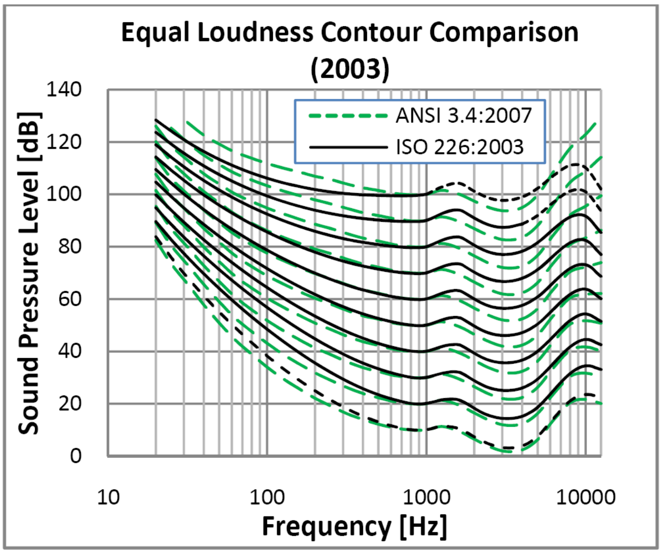

Figure 2 – Comparison of equal loudness contours as obtained from ANSI S3.4:2007 with ISO 226:2003 curves. ([7])

Also, the argument of more ISO 226 consistency was further doubted by new insight that loudness models matching previous ISO 226:1987 contours often give better agreement with subjective impressions than loudness models matching more recent ISO 226:2003 contours ([6],[9]).

For these reasons the argument of better matching to ISO 226 contours was highly controversial when being used to justify a general preference of one method with respect to the full spectrum of applicable sounds.

To further illustrate the difference of the methods, fig 3 shows results for pink noise as obtained from the calculation procedures of DIN 45631:1991 and ANSI S3.4:2007 ([10]). By tendency, this results seems to be representative in stating that the Moore/Glasberg approach tends to higher loudness values for broadband sounds. Typically, this difference relates to level differences of up to 5 dB.

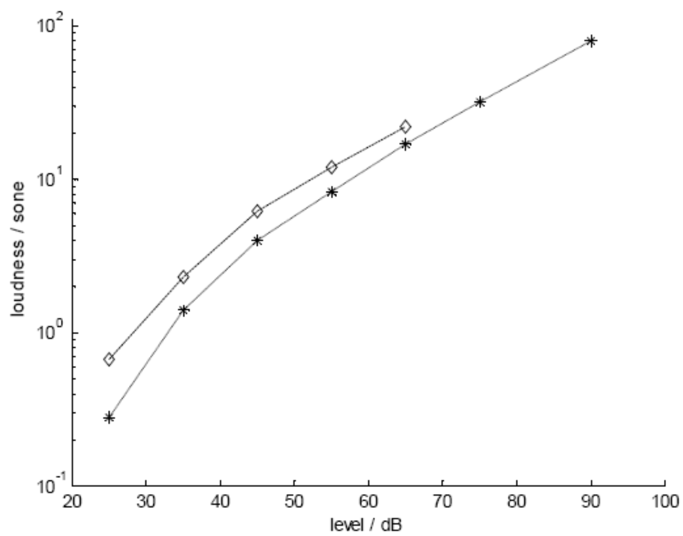

Figure 3 – Comparison of calculated loudness for pink noise as obtained from DIN 45631:1991 and ANSI S3.4:2007 ([10]).

In conclusion, it may be stated that the ANSI approach tends to higher loudness values for broadband sounds as well as to stronger suppressions of low frequency sound contributions. Although the results show unpredictable differences, these differences are found to be in the range of subjective assessment uncertainties.

5. SCIENTIFIC DEVELOPMENT AND NORMATIVE CONTINUITY

Standards – by their very nature – define a conflict area between the dynamical request of science to continuously follow any improvement or adaptation to new procedural insight and the static demand of industry for continuity to ensure widespread use and to allow for long-term comparison of sound characteristics. It is obvious that normative continuity is an important and legitimate requirement of the standard user community, crucially determining the benefit of a standard and thus representing a high value. This is because continuity only may guarantee long term comparability of results. This seems to be of particular importance for psychoacoustic metrics which did hard to get introduced and acknowledged after many years.

Therefore, particular care in changing standards is needed to ensure long term reliability and consistency in use against any desire for premature short term improvements. Significant improvements only justify the effort of essentially revising or modifying a well accepted and acknowledged standard. This is particularly true if this revision breaks off long term data continuity by introducing differing results for the same input data. Consequently, it was this requirement of continuity which has prevented the supporters of the Zwicker method to implement more than slight adaptations by now.

Of course, this legitimate interest of industry cannot be an excuse for inappropriate inflexibility against improved insight and scientific progress. But science really needs good and strong arguments to justify essential changes of established standards without risking or even loosing their credibility and convinced acceptance.

In spite of year long search for doubtless evidence of better results (with respect to better match with subjective impressions), no systematic, unique superiority of any approach could be found. The only convincing and traceable improvement with respect to “better” results could be proven for standardized (sinusoidal) hearing impressions as defined by ISO 226 equal loudness contours. However, the limited practical relevance of purely sinusoidal sounds together with serious doubts in better basing assessments of complex sounds on the match with new instead of previous ISO 226 curves prevented ISO TC 43 WG 9 from finding respective consensus.

6. RECENT DEVELOPMENTS AND OUTLOOK

Although being split and thus being far away from any consensus, a majority within ISO TC 43 WG 9 felt self-confident enough to remove the Zwicker method from being an ISO standard any more. However, the many passionate and intensive preceding discussions within WG 9 were taken up with high interest and commitment by the interested community, in particular by those having fixed long term evaluation and specification procedures on the basis of ISO 532 B. Missing any convincing evidence of the practical superiority of the new approach, many external institutions including science and industry expressed their strong concerns and objections against the tendency to remove ISO 532 B. Among these were the relevant suppliers of psychoacoustic evaluation software who expressed their strong opposition against any removal of a proven standard.

However, all such involvement was not able to put WG 9 off going on with exclusively prioritizing the new ANSI approach. Consequently, the initial decision to include both methods into new ISO 532 was given up by recommending the removal of the Zwicker method from ISO/DIS 532-1. Before adopting this recommendation, ISO TC 43 thoroughly discussed the many aspects of dissent and of the resulting normative discontinuity and finally expressed confidence that this controversial issue was going to be decided in a high sense of responsibility by the member bodies voting.

The discussions within TC 43 together with the many external involvements brought forward before now were taken up by intensive discussions between the member bodies of ISO TC 43 and interested institutions within their sphere of influence. Finally, these discussions ended up in a clear disapproval of ISO/DIS 532-1 by the member bodies vote together with the requirement to preserve the proven standard for the Zwicker method.

This requirement was picked up by ISO TC 43 WG 9 and resulted in preparing two separate standards,

- ISO 532-1 “Methods for calculating loudness – Part 1: Zwicker method” and
- ISO 532-2 “Methods for calculating loudness – Part 2: Moore/Glasberg method”

Meanwhile, respective committee drafts (ISO/CD) had been circulated to the ISO TC 43 member bodies for comments. Based on these comments, revised documents will be circulated as committee drafts again and it is planned to submit a formal draft of ISO 532 (ISO/DIS 532) by the end of the year.

ISO 532-1 will fully incorporate DIN 45631/A1 with full disclosure of all necessary implementation details including well documented software by source code. Thus, the implementation of the Zwicker method will be available in a traceable way for arbitrary non-stationary sounds including stationary sounds as a special case. By this, full continuity with previous implementations will be guaranteed.

By ISO 532-2 the Moore/Glasberg method will be available as a procedural alternative. Starting from an implementation of the approach for stationary sounds first, it is intended to extend the standard for non-stationary sounds later.

With these standards being introduced, the normative platform will be led back to widespread practice on an international level. This will help to further establish psychoacoustic metrics as an instrumental estimate of subjective hearing impressions. Of course it would be favourable to have better agreement between the results of each method. For the time being, unless further research results give new insight, the differences in numerical metrics must be handled by clear reference to the method used in each case.

7. SUMMARY

There are good reasons for a twin solution introducing two separate standards, ISO 532-1 and ISO 532-2, both defining alternative methods for calculation of loudness. Conceding that no practical superiority could be shown with respect to the multiplicity of real life sounds, this concept will maintain normative continuity by updating ISO 532 B along common practice only while following recent progress in replacing ISO 532 A by the new ANSI approach.

This opens the floor for comparing the approaches in practice. In the long term, this should allow that proven practical value and usefulness – together with new scientific insight – may offer a solid basis for establishing a unique standard later then.

REMARK

This paper was motivated by repeated requests to widely disseminate background information on the process and the related discussions and arguments which altogether finally led to the two separate standards as described above. The author gratefully acknowledges this widespread interest and is glad to present this information – after previous presentation on the European level (Euroregio 2013) – on the international level of Inter-Noise 2014 again. However, it should be mentioned that apart from some updating information on progress in the related ISO project, the content of this paper closely matches with [11].

REFERENCES

1. ISO 532: Acoustics – Method for calculating loudness level, 1975
2. DIN 45631: Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum - Verfahren nach Zwicker (Calculation of loudness level and loudness from the sound spectrum - Zwicker method), Berlin, Germany, 1991
3. DIN 45631/A1: Änderung 1: Berechnung der Lautheit zeitvarianter Geräusche – Amendment 1: Calculation of the loudness of time-variant sound), Berlin, Germany, 2010
4. ANSI S 3.4, Procedure for the computation of loudness of steady sounds, Melville, NY, USA, 2007
5. Hellman, R.: A new ANSI loudness standard. Acoustics today, ASA, January 2007
6. Scheuren, J.: Setting psychoacoustic standards and guidelines in the fields of interests between science and industry. Proc. Inter-Noise 2011, Osaka, 2011
7. Charbonneau, Novak, Ule: Comparison of loudness calculation procedure results to equal loudness contours. Proc. Inter-Noise 2009, Ottawa, Canada, 2009
8. Schomer, P.: Use of the new ISO 226 equal loudness contours as a filter to assess noise annoyance. Proc. Inter-Noise 2001, Den Haag, Netherlands, 2001
9. Schlittenlacher, J. et al: Loudness of pink noise and stationary technical sounds. Proc. Inter-Noise 2011, Osaka, 2011
10. Fastl, Völk, Straubinger: Standards for calculating loudness of stationary or time-variant sounds. Proc. Inter-Noise 2009, Ottawa, Canada, 2009
11. Scheuren, J.: Revision of ISO 532 "Acoustics - Methods for calculating loudness level" - recent development and future outline. AIA-DAGA/Euroregio 2013, Merano, Italy, 2013